

Convergence of an infinite dimensional stochastic process to a spatially structured trait substitution sequence

Hélène Leman*

September 6, 2017

Abstract

We consider an individual-based spatially structured population for Darwinian evolution in an asexual population. The individuals move randomly on a bounded continuous space according to a reflected brownian motion. The dynamics involves also a birth rate, a density-dependent logistic death rate and a probability of mutation at each birth event. We study the convergence of the microscopic process when the population size grows to $+\infty$ and the mutation probability decreases to 0. We prove a convergence towards a jump process that jumps in the infinite dimensional space of the stable spatial distributions. The proof requires specific studies of the microscopic model. First, we examine the large deviation principle around the deterministic large population limit of the microscopic process. Then, we find a lower bound on the exit time of a neighborhood of a stationary spatial distribution. Finally, we study the extinction time of the branching diffusion processes that approximate small size populations.

Keywords: structured population; birth and death diffusion process; large deviations studies; exit time; branching diffusion processes; nonlinear reaction diffusion equations; weak stability; Trait Substitution Sequence.

1 Introduction

The spatial aspect is an important issue in ecology [30, 12]. The influence of the heterogeneity of the environment on the phenotypic evolution has been explored for a long time [13, 17, 20]. For example, the emergence of phenotypic clusters under a heterogeneous space has been extensively studied [11, 27, 21]. In [11, 21], the authors suggest that clustering and aggregation of individuals can be a consequence of the spatial competition between individuals. Those phenomena generate a structured population based on isolated patches. In [27], the authors draw attention to the influence of the boundary of the spatial environment. The sensibility to heterogeneously distributed resources is also a key point to study the spatial dynamics of population [19]. In this context, the effect

*CMAP, Ecole Polytechnique, UMR 7641, route de Saclay, 91128 Palaiseau Cedex-France; E-mail: helene.leman@polytechnique.edu

of a spatial structure on the evolution of a population is fundamental.

In this paper, we use a population model that describes the interplay between evolution and spatial structure. We are interested in a macroscopic approximation of the microscopic model under three main biological assumptions : rare mutations, large population size and the impossibility of coexistence of two traits for a long time scale. The main result of this paper implies a convergence of the microscopic model to a jump process that jumps in an infinite dimensional space characterized by the spatial profiles of the population. This result is correlated with several works on adaptive dynamics and in particular with the model of Trait Substitution Sequence (TSS). Metz and al. [26] have introduced this model from an ecological point of view. It describes the evolution of phenotypic traits in the case where the ecological time scale, related to the dynamics of the population, and the evolutionary time scale, related to the mutations, are separated. In the evolution time scale, the model describes the succession of invading phenotypic traits as a jump Markov process in the space of phenotypic traits. The link between the microscopic model and the TSS model has been completely proved by Champagnat [3] in a simpler ecological context. Recently, some papers have generalized this approach in the case of an aged-structured population [31], of a multi-resources chemostat model [4] or of a prey-predator model [6]. But only [31] deals with some processes with values in infinite dimensional spaces, and the age structure is deterministic. In contrast with it, we are concerned with the spatial aspect of a population living on a heterogeneous environment where the individuals move randomly.

We study an individual-based model in which any individual birth and death events are described. This Markov process has been initially introduced by Champagnat and Méléard in [5]. For any time $t \geq 0$, each individual i is described by two characteristics (X_t^i, U_t^i) . X_t^i represents its location in an open, bounded and convex subset \mathcal{X} of \mathbb{R}^d with a C^2 -boundary. U_t^i denotes its phenotypic trait which belongs to a compact subset \mathcal{U} of \mathbb{R}^q . The phenotype of an individual does not change during its life time contrary to its location. The location space may represent a geographic landscape or a theoretical space that describes a gradient of temperature, a gradient of elevation or a resource parameter as seed size for a population of birds [11]. In the context of the last example, the phenotype may represent the beak size of a bird and it is interesting to study the adequacy between the beak size and the seed size when evolution occurs [19].

The total population is represented at any time t by the finite measure

$$\nu_t^K = \frac{1}{K} \sum_{i=1}^{N_t} \delta_{(X_t^i, U_t^i)}, \quad (1.1)$$

where δ_y corresponds to the Dirac measure at y , N_t is the number of individuals at time t . The parameter K scales the population size and the biological assumption of large population size is stated into mathematics by K tends to $+\infty$.

The dynamics of the process is driven by a birth and death diffusion process, in which the motion, birth, mutation and death of each individual depends on its location and trait.

Any individual i with phenotypic trait u moves according to a diffusion process driven

by the following stochastic differential equation normally reflected at the boundary $\partial\mathcal{X}$,

$$dX_t^i = \sqrt{2m^u} Id \cdot dB_t - n(X_t^i)dl_t \quad (1.2)$$

where B is a d -dimensional brownian motion, l_t is an adapted continuous and non-decreasing process with $l_0 = 0$, it increases only when $X_t^i \in \partial\mathcal{X}$ and the diffusion coefficient m^u is a function of the trait.

We consider a population with asexual reproduction. An individual with location $x \in \mathcal{X}$ and trait $u \in \mathcal{U}$ gives birth at rate $b(x, u)$. This rate can also be denoted by $b^u(x)$ when u is fixed and it is assumed to be bounded. The offspring appears at the location of its parents. Furthermore, a mutation may occur with probability $q_K p$, making the phenotypic trait of the offspring different. The law of the mutant trait is then given by a kernel $k(x, u, \cdot)$. The mutation probability p may depend on the trait and the location. The parameter q_K scales the mutation probability and the biological assumption of rare mutations is stated by $q_K \rightarrow 0$.

The death rate depends on the characteristics of the individual and on the competition between all individuals. The natural death rate is $d(x, u)$. The competition exerted by an individual (y, v) on an individual (x, u) depends on the location y and on the two traits through a competition kernel $c : \mathcal{U} \times \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R}^+$. For the population $\nu = \frac{1}{K} \sum_{i=1}^n \delta_{(x_i, u_i)} \in M_F(\mathcal{X} \times \mathcal{U})$, the competitive pressure exerted on individual (x, u) is

$$c \cdot \nu(x, u) = \frac{1}{K} \sum_{i=1}^n c(u, x_i, u_i) = \frac{1}{K} \int_{\mathcal{X} \times \mathcal{U}} c(u, y, v) \nu(dy, dv).$$

Remark that the competition kernel does depend on y . This spatial dependence yields non-trivial mathematical difficulties. In Champagnat-Méléard [5], the competition kernel depends also on x but the long time behavior of the deterministic limit is still unknown, to our knowledge. Finally, the total death rate is $d(x, u) + c \cdot \nu(x, u)$. As for the birth rate, $d(x, u)$ can also be denoted by $d^u(x)$ and $c(u, y, v)$ by $c^{uv}(y)$ when u and v are fixed. Let us state the assumptions on the parameters.

- Assumption 1.1.**
1. m, b, d, k and c are continuous and non-negative on their domains and b, d and c are Lipschitz functions with respect to x and y .
 2. There exist $\bar{m}, \underline{b}, \bar{b}, \bar{d}, \bar{c}, \underline{c}, \bar{k} \in \mathbb{R}$ such that for any $(x, u, y, v) \in (\bar{\mathcal{X}} \times \mathcal{U})^2$, $0 < m^u \leq \bar{m}$, $\underline{b} < b(x, u) \leq \bar{b}$, $d(x, u) \leq \bar{d}$, $\underline{c} \leq c(u, y, v) \leq \bar{c}$, $k(x, u, v) \leq \bar{k}$, and d is not the zero function.
 3. The sequence of initial measures $(\nu_0^K)_{K>0}$, which belongs to $M_F(\mathcal{X} \times \mathcal{U})$, converges in law to some deterministic measure denoted by ξ_0 and it satisfies $\sup_K \mathbb{E}[\langle \nu_0^K, 1 \rangle^3] < +\infty$.
 4. q_K tends to 0 when K tends to $+\infty$.

Before going further, let us set and recall the notation, which we use in the entire paper.

Notation

- For all $x \in \partial\mathcal{X}$, $n(x)$ denotes the outward normal to the boundary of \mathcal{X} at point x .
- For sufficiently smooth f and for all $(x, u) \in \partial\mathcal{X} \times \mathcal{U}$, $\partial_n f(x, u)$ denotes the scalar product $\nabla_x f(x, u) \cdot n(x)$.
- $C_n^{k,l}(\mathcal{X} \times \mathcal{U})$ represents the set of functions f such that $f \in C^{k,l}(\mathcal{X} \times \mathcal{U})$ and $\partial_n f(x, u) = 0$ for all $(x, u) \in \partial\bar{\mathcal{X}} \times \mathcal{U}$. We define $C_n^{k,l,j}(\mathcal{X} \times \mathcal{U} \times [0, T])$ similarly.
- For any $f \in C_n^{k,l,j}(\mathcal{X} \times \mathcal{U} \times [0, T])$, f_s is the function on $\mathcal{X} \times \mathcal{U}$ such that $f_s(x, u) = f(x, u, s)$.
- For any compact set \mathbb{X} , we denote the space of finite measures on \mathbb{X} by $M_F(\mathbb{X})$.
- $C^{Lip}(\mathbb{X})$ denotes the set of all positive Lipschitz-continuous functions f on \mathbb{X} bounded by 1 and with a Lipschitz constant smaller than 1.
- We define the Kantorovich-Rubinstein distance on $M_F(\mathbb{X})$ by : for any $\nu, \mu \in M_F(\mathbb{X})$,

$$\mathcal{W}_1(\nu, \mu) = \sup_{f \in C^{Lip}(\mathbb{X})} \left| \int_{\mathbb{X}} f d\nu - \int_{\mathbb{X}} f d\mu \right|.$$

As \mathbb{X} is a compact set, this metric is a metrization of the topology of weak convergence. It is equivalent to the 1st-Wasserstein distance.

- $B(\nu, \gamma)$ represents the ball of center ν and radius γ in $M_F(\mathbb{X})$ for the previous distance.
- $\mathbb{D}([0, T], M_F(\mathbb{X}))$ denote the space of càdlàg functions from $[0, T]$ to $M_F(\mathbb{X})$, equipped with the Skorokhod topology.
- For any $\xi \in M_F(\mathcal{X} \times \{u, v\})$, we identify the two following ways of writing : $\xi(dx, dw) = \xi^u(dx)\delta_u(dw) + \xi^v(dx)\delta_v(dw)$ and $\xi = (\xi^u, \xi^v) \in (M_F(\mathcal{X}))^2$.

2 Main theorem

A full algorithmic description and a mathematical formulation of the model described in the previous part are detailed in Champagnat and Méléard [5]. Moreover, a macroscopic approximation has been proved as a large population limit.

Theorem 2.1 (Theorems 4.2 and 4.6 in [5]). *Suppose that Assumption 1.1 holds. For all $T > 0$, the sequence $(\nu^K)_{K>0}$ of processes belonging to $\mathbb{D}([0, T], M_F(\mathcal{X} \times \mathcal{U}))$ converges in law to a deterministic and continuous function ξ , i.e. $\xi \in \mathbb{C}([0, T], M_F(\mathcal{X} \times \mathcal{U}))$ such that $\sup_{t \in [0, T]} \langle \xi_t, 1 \rangle < +\infty$ and $\forall f \in C_n^{2,0}(\mathcal{X} \times \mathcal{U})$,*

$$\langle \xi_t, f \rangle = \langle \xi_0, f \rangle + \int_0^t \int_{\mathcal{X} \times \mathcal{U}} \left\{ m^u \Delta_x f(x, u) + [b(x, u) - d(x, u) - c \cdot \xi_s(x, u)] f(x, u) \right\} \xi_s(dx, du) ds. \quad (2.1)$$

Moreover, if \mathcal{U} is finite, for any $u \in \mathcal{U}$ and $t > 0$, $\xi_t(., u)$ has a density with respect to Lebesgue measure which is a C^2 -function.

The limiting equation (2.1) is a nonlinear nonlocal reaction-diffusion equation defined on the space of traits and locations. In [10, 1], the authors have studied the existence of

the steady states of similar equations in the context of frequent mutations. Our study involves a rare mutations assumption and mutation terms disappear in the limit. The stability of the steady states and the long time behavior of the solutions to (2.1) have been characterized in [22, 7] in the particular cases of a monomorphic population (all individuals have a same phenotype) and a dimorphic population (two traits are involved). The stationary states and their stability are described using the following parameters.

Definition 2.2. For any $u \in \mathcal{U}$, we define H^u by,

$$H^u = - \min_{\phi \in H^1(\mathcal{X}), \phi \neq 0} \frac{1}{\|\phi\|_{L^2(\mathcal{X})}^2} \left[m^u \int_{\mathcal{X}} |\nabla \phi(x)|^2 dx - \int_{\mathcal{X}} (b^u - d^u)(x) \phi^2(x) dx \right], \quad (2.2)$$

where $H^1(\mathcal{X})$ is the Sobolev space of order 1 on \mathcal{X} . H^u is thus the principal eigenvalue of the operator $m^u \Delta_x \cdot + (b^u - d^u) \cdot$ with Neumann boundary condition on \mathcal{X} . Let $\bar{g}^u \in H^1(\mathcal{X})$ be the eigenfunction of the previous operator associated with the eigenvalue H^u such that

$$\int_{\mathcal{X}} c^{uu}(y) \bar{g}^u(y) dy = H^u.$$

According to [22], $\bar{g}^u \in C^1(\mathcal{X})$. If $\bar{g}^u \geq 0$, we define the associated measure in $M_F(\mathcal{X})$

$$\bar{\xi}^u(dx) := \bar{g}^u(x) dx.$$

Finally, for any $(u, v) \in \mathcal{U}$, we set

$$\kappa^{vu} := \int_{\mathcal{X}} c^{vu}(y) \bar{g}^u(y) dy \left(\int_{\mathcal{X}} \bar{g}^u(y) dy \right)^{-1}.$$

As proved in [22], $H^u > 0$ is the condition ensuring that a monomorphic population with trait u is able to survive for a long time at the ecological time scale. In that case, the stationary stable state is described by the positive spatial profile \bar{g}^u . The dimorphic case implies four distinct stationary states : the trivial state $(0, 0)$, two monomorphic states and one co-existence state. To ensure that the co-existence state is unstable, we set the following assumption.

Assumption 2.3. Let u be in \mathcal{U} , for almost all $v \in \mathcal{U}$,

1. either, $H^v \kappa^{uu} - H^u \kappa^{vu} < 0$,
2. or, $\begin{cases} H^v \kappa^{uu} - H^u \kappa^{vu} > 0 \\ H^u \kappa^{vv} - H^v \kappa^{uv} < 0. \end{cases}$

Using [22], we notice that Assumption 2.3 states into mathematics the impossibility of co-existence of two traits for a long time, this assumption is also known as the "Invasion-Implies-Fixation" principle. Under Assumption 2.3, any solution to (2.1) converges either to $(\bar{g}^u, 0)$, or to $(0, \bar{g}^v)$, which are two monomorphic states. More precisely,

Condition 1 ensures the stability of the equilibrium $(\bar{g}^u, 0)$. Thus, if a mutant population with phenotype v is emerging in a monomorphic well-established population with phenotype u , it will not be able to survive. Under Condition 2, the authors of [22] prove that the deterministic solution to (2.1) converges to the stable equilibrium $(0, \bar{g}^v)$ whatever the initial condition is. In the light of the previous considerations, we refer to $H^v \kappa^{uu} - H^u \kappa^{vu}$ as the invasion fitness of the individuals with type v in a resident population with type u . Furthermore, the probability of success of such an invasion is described precisely by means of the geographical birth position x_0 of the first individual with trait v and the function ϕ^{vu} defined below. That probability is precisely $\phi^{vu}(x_0)$.

Definition 2.4. For any $u, v \in \mathcal{U}$, ϕ^{vu} is the function on \mathcal{X} such that

1. If $H^v \kappa^{uu} - H^u \kappa^{vu} \leq 0$, $\phi^{vu}(x) = 0$ for all $x \in \mathcal{X}$.
2. If $H^v \kappa^{uu} - H^u \kappa^{vu} > 0$, ϕ^{vu} is the unique positive solution to the elliptic equation
$$\begin{cases} m^v \Delta_x \phi(x) + \left(b^v(x) - d^v(x) - \int_{\mathcal{X}} c^{vu}(y) \bar{g}^u(y) dy \right) \phi(x) - b^v(x) \phi(x)^2 = 0, \forall x \in \mathcal{X}, \\ \partial_n \phi(x) = 0, \forall x \in \partial \mathcal{X}. \end{cases} \quad (2.3)$$

We are now ready to state the main result of this paper.

Theorem 2.5. We suppose that Assumptions 1.1 and 2.3 hold. We also assume that the scaling parameters satisfy

$$K q_K \log(K) \xrightarrow{K \rightarrow +\infty} +\infty \quad \text{and} \quad K q_K e^{KV} \xrightarrow{K \rightarrow +\infty} 0, \quad \text{for any } V > 0. \quad (2.4)$$

Then for any $T > 0$, $\left(\nu_{(t/Kq_K)}^K \right)_{t \in [0, T]}$ converges towards a jump Markov process $(\Lambda_t)_{t \geq [0, T]}$ as $K \rightarrow +\infty$. At any time t , Λ_t belongs to the subspace $\{\bar{\xi}^u \delta_u, u \in \mathcal{U}\}$ of $M_F(\mathcal{X} \times \mathcal{U})$, where for any $u \in \mathcal{U}$, $\bar{\xi}^u \in M_F(\mathcal{X})$ is the spatial pattern defined in Definition 2.2. The process $(\Lambda_t)_{t \geq 0}$ jumps from the state characterized by the trait $u \in \mathcal{U}$ to the state characterized by $v \in \mathcal{U}$ at the infinitesimal rate

$$\int_{\mathcal{X}} p b^u(x) \phi^{vu}(x) \bar{g}^u(x) k(x, u, v) dx dv.$$

This convergence holds in the sense of convergence of the finite dimensional distributions.

Remark that $\bar{\xi}^u$ describes the spatial distribution of the monomorphic population with trait u . The limiting process jumps from a spatial distribution to another one depending on the mutant trait. It models an evolutionary phenomenon using a sequence of monomorphic equilibria described by their spatial patterns.

Although the structure of Theorem 2.5's proof is similar to the one of Theorem 1 in [3], the spatial structure of the process leads us to deal with infinite dimensional processes. Two key points of the proof have to be approached differently. The first point

concerns the study of the process when it is close to a monomorphic deterministic equilibrium. The aim is to estimate the exit time of a neighborhood of a stationary state to (2.1). We give the behavior of the stochastic process around its deterministic equilibrium taking into account a small mutant population and the possibility of other mutations. Thanks to it, we avoid the comparisons used in [3], where the behavior of the resident population process is compared with the behavior of a theoretical monomorphic population evolving alone. Those comparisons are much more involved when the population is spatially structured. Moreover, to estimate this exit time, we have to study a large deviation principle of the stochastic process $(\nu_t^K)_{t \geq 0}$ around its deterministic limit (2.1) when K is large. The large deviations studies for processes combining diffusion process and jumps are still unresolved, to our knowledge. Those studies have thus their own interest. The second point which is approached differently concerns the study of small population size processes. The aim is to understand the dynamics of a population descended from a mutant which has appeared in a well-established monomorphic population. As long as the mutant population size is small, the competitive terms between mutants can be neglected. Thus, the dynamics of the mutant population can be compared with the dynamics of a branching diffusion process. We describe finely the survival probability of a branching diffusion process by means of the eigenparameters defined previously and we link it with the conditions presented in Assumption 2.3.

In Section 3, we explicit the upper bound of the large deviation principle by using ideas in [9, 25, 31]. Then we study the functional rate associated with the large deviation principle. Section 4 deals with the exit time of a neighborhood of a stationary state to (2.1). In Section 5, we study a branching diffusion process. First, we evaluate its probability of survival. Then, we characterize the scale time under which its size is of order K . Section 6 is devoted to the proof of Theorem 2.5. We detail two key propositions. The first one deals with the dynamics of the individual-based process in the case where only one trait are involved. The second one gives the dynamics of the process after the time of the first mutation but as long as only at most two traits are involved. Finally, Section 7 presents a numerical example that illustrates Theorem 2.5.

3 Exponential deviations results

In this section, we are concerned by the large deviations from the large population limit (2.1) for the process $(\nu_t^K)_{t \in [0, T]}$ when K tends to $+\infty$ and q_K tends to 0. First of all, Theorem 3.1 gives the upper bound of the large deviations principle. This theorem involves a rate function. Let us first explicit it. That requires specific notation which will be only used in this subsection : let us fix $T > 0$,

- $\mathcal{E} = \mathcal{X} \times \mathcal{U} \times \{1, 2\}$.
- ψ is the mapping such that for any function $f \in C^{2,0,1}(\mathcal{X} \times \mathcal{U} \times [0, T])$, for any

$$(x, u, \pi, t) \in \mathcal{E} \times [0, T],$$

$$\psi(f)(x, u, t, \pi) = \begin{cases} f(x, u, t) & \text{if } \pi = 1, \\ -f(x, u, t) & \text{if } \pi = 2. \end{cases}$$

- For all $\nu = (\nu_t)_{t \in [0, T]} \in \mathbb{D}([0, T], M_F(\mathcal{X} \times \mathcal{U}))$, we define the positive finite measure

$$\mu_t^\nu(dx, du, d\pi) = [b(x, u)\delta_1(d\pi) + (d(x, u) + c \cdot \nu_{t-}(x, u))\delta_2(d\pi)]\nu_{t-}(dx, du). \quad (3.1)$$

- Finally, we introduce the log-Laplace transform ρ of a centered Poisson distribution with parameter 1, $\rho(x) = e^x - x - 1$, and its Legendre transform ρ^* ,

$$\rho^*(y) = ((y+1)\log(y+1) - y)\mathbb{1}_{\{y > -1\}} + \mathbb{1}_{\{y = -1\}} + \infty \cdot \mathbb{1}_{\{y < -1\}}.$$

We are now ready to define the rate function in which we shall be interested : for all $\xi_0 \in M_F(\mathcal{X} \times \mathcal{U})$ and $\nu \in \mathbb{D}([0, T], M_F(\mathcal{X} \times \mathcal{U}))$,

$$I_{\xi_0}^T(\nu) := \begin{cases} \sup_{f \in C_n^{2,0,1}(\mathcal{X} \times \mathcal{U} \times [0, T])} I^{f,T}(\nu), & \text{if } \nu_0 = \xi_0 \\ +\infty, & \text{otherwise,} \end{cases} \quad (3.2)$$

where

$$\begin{aligned} I^{f,T}(\nu) := & \langle \nu_T, f_T \rangle - \langle \nu_0, f_0 \rangle - \int_0^T \langle m\Delta_x f_s + m|\nabla_x f_s|^2 + \frac{\partial f_s}{\partial s}, \nu_s \rangle ds \\ & - \int_0^T \int_{\mathcal{E}} \left(\psi(f)(x, u, s, \pi) + \rho(\psi(f)(x, u, s, \pi)) \right) d\mu_s^\nu ds. \end{aligned}$$

When there is no ambiguity, we will write $I^T(\nu)$ instead of $I_{\nu_0}^T(\nu)$.

Theorem 3.1. *Suppose that Assumptions 1.1 holds. For all $\alpha > 0$, $\xi_0 \in M_F(\mathcal{X} \times \mathcal{U})$, for all compact set $\mathcal{C} \subset B(\xi_0, \alpha)$, for all measurable subset A of $\mathbb{D}([0, T], M_F(\mathcal{X} \times \mathcal{U}))$ such that there exists $M > 0$ with $A \subset \{\nu | \sup_{t \in [0, T]} \langle \nu_t, \mathbb{1} \rangle \leq M\}$,*

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \sup_{\nu_0^K \in \mathcal{C} \cap M_F^K} \log \mathbb{P}_{\nu_0^K}(\nu^K \in A) \leq - \inf_{\xi \in \mathcal{C}, \nu \in \bar{A}} I_\xi^T(\nu), \quad (3.3)$$

where $M_F^K = \{\frac{1}{K} \sum_{i=1}^N \delta_{(x_i, u_i)}, \text{ with } N \in \mathbb{N}, (x_i, u_i) \in \mathcal{X} \times \mathcal{U}\}$.

Proof. We will show the following upper bound

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \mathbb{P}(\nu^K \in A) \leq - \inf_{\nu \in \bar{A}} I_{\xi_0}^T(\nu), \quad (3.4)$$

indeed (3.3) can be directly deduced from this bound by a similar reasoning as in the proof of Corollary 5.6.15 in Dembo and Zeitouni [9]. To prove (3.4), we need the exponential tightness of the process $(\nu_t^K)_{t \in [0, T]}$ which is described by the following lemma.

Lemma 3.2. *Suppose that Assumption 1.1 holds, and that there exists $C_{init} > 0$ such that $\sup_{K \in \mathbb{N}} \langle \nu_0^K, \mathbb{1} \rangle < C_{init}$ a.s.. Then for all $L > 0$, there exists a compact subset \mathcal{C}_L of the Skorohod space $\mathbb{D}([0, T], M_F(\mathcal{X} \times \mathcal{U}))$ such that*

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \mathbb{P}(\nu^K \notin \mathcal{C}_L) \leq -L$$

We do not detail the proof of Lemma 3.2 as it may be easily adapted from [8, 18, 31]. Then, set $\tau_M^K = \inf\{t \geq 0, \langle \nu_t^K, \mathbb{1} \rangle \geq M\}$. Note that Lemma 3.2 is also true for $(\nu_{t \wedge \tau_M^K}^K)_{t \geq 0}$. Using a proof similar to Theorem 4.4.2 of [9], we deduce the inequality

$$\begin{aligned} \limsup_{K \rightarrow +\infty} \frac{1}{K} \log \mathbb{P}(\nu^K \in A) &= \limsup_{K \rightarrow +\infty} \frac{1}{K} \log \mathbb{P}(\nu_{\cdot \wedge \tau_M^K}^K \in A) \\ &\leq - \inf_{\nu \in \bar{A}} \left(\sup_{f \in C_n^{2,0,1}(\mathcal{X} \times \mathcal{U} \times [0, T])} (I^{f,T}(\nu) - H(I^{f,T})) \right), \end{aligned}$$

where $H(I^{f,T}) = \limsup_{K \rightarrow +\infty} \frac{1}{K} \log \mathbb{E}[\exp(K I^{f,T}(\nu_{\cdot \wedge \tau_M^K}^K))]$. Let us show that $H(I^{f,T}) = 0$. Let

$$\begin{aligned} \mathcal{N}_T &= \exp \left(K I^{f,T}(\nu_{\cdot \wedge \tau_M^K}^K) \right. \\ &\quad \left. - q_K K \int_0^{T \wedge \tau_M^K} \left\langle \nu_s^K, p b^u(x) \left(\int_{\mathcal{U}} \phi(f_s)(x, w) k(x, u, w) dw - \phi(f_s)(x, u) \right) \right\rangle ds \right), \end{aligned}$$

where $\phi(x) = x + \rho(x) = e^x - 1$. Itô's Formula implies that $(\mathcal{N}_T, T \geq 0)$ is a local martingale. The definition of τ_M^K implies that \mathcal{N}_T is bounded. So it is a martingale of mean 1 and there exists a constant $C(\|f\|_\infty, M) > 0$ such that

$$\exp(-q_K K C(\|f\|_\infty, M)) \leq \mathbb{E} \left[\exp \left(K I^{f,T}(\nu_{\cdot \wedge \tau_M^K}^K) \right) \right] \leq \exp(q_K K C(\|f\|_\infty, M)).$$

We conclude easily, since q_K tends to 0 when $K \rightarrow +\infty$. \square

The aim is now to write the rate function under a non-variational integral formulation which is more workable than that of (3.2). Firstly, this integral formulation is convenient to use Chasles' Theorem. Secondly, it will be used to bound from above the distance between a solution to (2.1) and any ν , this upper bound is proved below in Proposition 3.4. Those two points are required to prove the results about the exit time in Section 4. Before writing the non-variational formulation, let us define two functional spaces.

- The Orlicz space associated with ρ^* is $L_{\rho^*, T}$ the set of all bounded and measurable functions h on $\mathcal{E} \times [0, T]$ such that

$$\|h\|_{\rho^*, T} := \inf \left\{ \alpha > 0, \int_{\mathcal{E} \times [0, T]} \rho^* \left(\frac{|h|}{\alpha} \right) d\mu_s^\nu ds \leq 1 \right\} < +\infty. \quad (3.5)$$

The Orlicz space associated with ρ is defined on the same way.

- \mathcal{L}_T^2 is the set of functions $h \in L^2(\mathcal{X} \times \mathcal{U} \times [0, T], \mathbb{R}^d)$ such that

$$\|h\|_{\mathcal{L}^2, T} := \left(\int_0^T 2\langle \nu_s, m|h_s|^2 \rangle ds \right)^{1/2} < \infty. \quad (3.6)$$

Theorem 3.3. *Suppose that Assumption 1.1 holds. Let $T > 0$ and $\nu \in \mathbb{D}([0, T], M_F(\mathcal{X} \times \mathcal{U}))$, such that $I_{\nu_0}^T(\nu) < +\infty$, then there exist two measurable functions $(h_1^\nu, h_2^\nu) \in L_{\rho^*, T} \times \mathcal{L}_T^2$ such that for all $f \in C_n^{2,0,1}(\mathcal{X} \times \mathcal{U} \times [0, T])$,*

$$\begin{aligned} \langle \nu_t, f_t \rangle &= \langle \nu_0, f_0 \rangle + \int_0^T \int_{\mathcal{E}} (1 + h_1^\nu(x, u, s, \pi)) \psi(f)(x, u, s, \pi) d\mu_s^\nu ds \\ &\quad + \int_0^T \int_{\mathcal{X} \times \mathcal{U}} \left(m^u \Delta_x f_s(x, u) + 2m^u h_2^\nu(x, u, s) \cdot \nabla_x f_s(x, u) + \frac{\partial f_s}{\partial s}(x, u) \right) \nu_s(dx, du), \end{aligned} \quad (3.7)$$

and the rate function can be written as follows

$$I_{\nu_0}^T(\nu) = \int_0^T \int_{\mathcal{E}} \rho^*(h_1^\nu) d\mu_s^\nu ds + \int_0^T m \langle \nu_s, |h_2^\nu|^2 \rangle ds < +\infty. \quad (3.8)$$

The proof of Theorem 3.3 uses convex analysis arguments which can be adapted from Leonard [23, 24]. We do not detail its proof but we give the main ideas. For all $\nu \in \mathbb{D}([0, T], M_F(\mathcal{X} \times \mathcal{U}))$, $I^T(\nu)$ is equal to the Legendre transform Γ^* of

$$\Gamma : \begin{pmatrix} (\psi, \nabla_x)(C^{2,0,1}(\mathcal{X} \times \mathcal{U} \times [0, T])) \\ (g_1, g_2) \end{pmatrix} \begin{matrix} \rightarrow \mathbb{R} \\ \mapsto \int_0^T \int_{\mathcal{E}} \rho(g_1) d\mu_s^\nu ds + \int_0^T \langle \nu_s, m|g_2|^2 \rangle ds \end{matrix} \end{pmatrix}. \quad (3.9)$$

at a well chosen point l_ν . If l_ν belongs to the interior of the set $\text{dom} \Gamma^*$ of linear maps l with $\Gamma^*(l) < +\infty$, we can exhibit l_ν by means of the derivative of $\bar{\Gamma}$, the Legendre biconjugate of Γ . Studying directly $\bar{\Gamma}$ is difficult. The key point is thus to work on the product space $L_{\rho, T} \times \mathcal{L}_T^2$. In this way, we can study the Legendre biconjugate of an extension of Γ on that space, in order to deal with the diffusive part and the jumps part separately. The diffusive part is treated using ideas of Dawson and Gartner [8] and Fontbona [14] whereas the jumps part is treated using ideas of Leonard [23, 24]. The next step is to deduce the Legendre biconjugate of Γ by restricting the definition domain. Finally, to deal with points ν for which l_ν does not belong to the interior of $\text{dom} \Gamma^*$, we use a continuity argument similar to that of Theorem 7.1's proof in [25].

The last result of this part gives an upper bound on the distance between a solution to (2.1) and any ν , this bound is used in Subsection 4.2.

Proposition 3.4. *Let $T > 0$ and $M > 0$. There exists $C(T, M)$ such that, for any ν satisfying $\sup_{t \leq T} \langle \nu_t, \mathbb{1} \rangle < M$ and for all $(\xi_t)_{t \geq 0}$ solution to (2.1) with the initial condition $\xi_0 = \nu_0$,*

$$\sup_{t \in [0, T]} \mathcal{W}_1(\nu_t, \xi_t) \leq C(T, M) \left(I^T(\nu) + \sqrt{I^T(\nu)} \right).$$

Proof. Let ν be such that $\langle \nu_t, \mathbb{1} \rangle < M$ for all $t \in [0, T]$. If $I^T(\nu) = 0$, i.e. $\nu_t = \xi_t$, or if $I^T(\nu) = +\infty$, the result is obvious, so let us assume that $0 < I^T(\nu) < +\infty$. Theorem 3.3 implies the existence of $(h_1, h_2) \in L_{\rho^*, T} \times \mathcal{L}_T^2$ such that

$$I^T(\nu) = \int_0^T \int_{\mathcal{E}} \rho^*(h_1) d\mu_s^\nu ds + \int_0^T m \langle \nu_s, |h_2|^2 \rangle ds. \quad (3.10)$$

We easily deduce that for any $t \leq T$,

$$\|h_2\|_{\mathcal{L}^2, t}^2 \leq \|h_2\|_{\mathcal{L}^2, T}^2 \leq 2I^T(\nu). \quad (3.11)$$

Let us also find an upper bound on $\|h_1\|_{\rho^*, T}$. Note that for all $x \in \mathbb{R}$,

$$\begin{cases} \text{if } \alpha \geq 1, & \rho^*(|x|/\alpha) \leq \rho^*(|x|)/\alpha \leq \rho^*(x)/\alpha, \\ \text{if } 0 < \alpha \leq 1, & \rho^*(|x|/\alpha) \leq \rho^*(|x|)/\alpha^2 \leq \rho^*(x)/\alpha^2. \end{cases} \quad (3.12)$$

Moreover, the non-variational formulation (3.10) implies that $I^T(\nu) \geq \int_0^T \int_{\mathcal{E}} \rho^*(h_1) d\mu_s^\nu ds$. Thus, using (3.12) and the definition of the norm $\|\cdot\|_{\rho^*, T}$ in (3.5), we obtain that if $I^T(\nu) \geq 1$, $\int_0^T \int_{\mathcal{E}} \rho^*(|h_1|/I^T(\nu)) d\mu_s^\nu ds \leq 1$, i.e. $\|h_1\|_{\rho^*, T} \leq I^T(\nu)$, and if $I^T(\nu) \leq 1$, $\int_0^T \int_{\mathcal{E}} \rho^*(|h_1|/\sqrt{I^T(\nu)}) d\mu_s^\nu ds \leq 1$, i.e. $\|h_1\|_{\rho^*, T} \leq \sqrt{I^T(\nu)}$. Thus, for any $t \leq T$,

$$\|h_1\|_{\rho^*, t} \leq \|h_1\|_{\rho^*, T} \leq \left(I^T(\nu) + \sqrt{I^T(\nu)} \right). \quad (3.13)$$

Let $(\xi_t)_{t \geq 0}$ be the solution to (2.1) with initial condition ν_0 . We want now evaluate $\mathcal{W}_1(\nu_t, \xi_t)$. Let us denote the semigroup of the reflected diffusion process which is the solution to (1.2) with the initial condition x and the diffusion coefficient m^u by $(P_t^u)_{t \geq 0}$. Using Theorem 3.3, we find the following mild formulation for $(\nu_t)_{t \geq 0}$ in a similar way to Lemma 4.5 in [5] : for all $f \in C^{Lip}(\mathcal{X} \times \mathcal{U})$,

$$\langle \nu_t, f \rangle = \int_0^t \int_{\mathcal{E}} \psi(P_{t-s} f) d\mu_s^\nu ds + \int_0^t \int_{\mathcal{E}} \psi(P_{t-s} f) h_1 d\mu_s^\nu ds + \int_0^t \langle \nu_s, 2m \nabla_x P_{t-s} f \cdot h_2 \rangle ds. \quad (3.14)$$

In addition with a mild equation for $(\xi_t)_{t \geq 0}$, we deduce that for all $f \in C^{Lip}(\mathcal{X} \times \mathcal{U})$ and for all $t \leq T$,

$$\begin{aligned} |\langle \nu_t - \xi_t, f \rangle| &= \left| \int_0^t \int_{\mathcal{E}} \psi(P_{t-s} f) (d\mu_s^\nu ds - d\mu_s^\xi ds) + \int_0^t \int_{\mathcal{E}} \psi(P_{t-s} f) h_1 d\mu_s^\nu ds \right. \\ &\quad \left. + \int_0^t \langle \nu_s, 2m \nabla_x P_{t-s} f \cdot h_2 \rangle ds \right| \\ &\leq C_1 \int_0^t \sup_{r \in [0, s]} \mathcal{W}_1(\nu_r, \xi_r) ds + \|\psi(P_{t-} f)\|_{\rho, t} \|h_1\|_{\rho^*, t} + \|\nabla_x P_{t-} f\|_{\mathcal{L}^2, t} \|h_2\|_{\mathcal{L}^2, t}. \end{aligned} \quad (3.15)$$

The second line is a consequence of Hölder's inequality (see for example Theorem 6 of Chapter 1 in [28] about Hölder's inequalities in Orlicz spaces).

Furthermore, the following Lemma insures that $P_t^u f \in C^{Lip}(\mathcal{X})$.

Lemma 3.5 (Part 2 of [32]). *As \mathcal{X} is a convex set in \mathbb{R}^d , for all $f \in C^{Lip}(\mathcal{X})$, $u \in \mathcal{U}$, and $t \in \mathbb{R}^+$, $P_t^u f \in C^{Lip}(\mathcal{X})$.*

Let us now find an upper bound on $\|\psi(P_{t-}f)\|_{\rho,t}$. f belongs to $C^{Lip}(\mathcal{X} \times \mathcal{U})$ and $\sup_{t \in [0,T]} \langle \nu_t, \mathbb{1} \rangle \leq M$. So, for all $\alpha > 0$, $t \leq T$

$$\left| \int_0^t \int_{\mathcal{E}} \rho \left(\frac{|\psi(P_{t-s}f)|}{\alpha} \right) d\mu_s^\nu ds \right| \leq \left| \int_0^T \int_{\mathcal{E}} \rho \left(\frac{1}{\alpha} \right) d\mu_s^\nu ds \right| \leq TM[\bar{b} + \bar{d} + \bar{c}M] \rho \left(\frac{1}{\alpha} \right),$$

so, for any $t \leq T$,

$$\|\psi(P_{t-}f)\|_{\rho,t} \leq \left[(\rho|_{\mathbb{R}^+})^{-1} \left(\frac{1}{TM[\bar{b} + \bar{d} + \bar{c}M]} \right) \right]^{-1} := C_2.$$

Furthermore, Lemma 3.5 implies also that for any $t \leq T$,

$$\|\nabla_x P_{t-}f\|_{\mathcal{L}^2,t}^2 = \int_0^t 2\langle \nu_s, m|\nabla_x P_{t-s}f|^2 \rangle ds \leq 2\bar{m}Mt \leq 2\bar{m}MT := C_3.$$

Using the last two inequalities with (3.11), (3.13) and (3.15), we find

$$\sup_{r \in [0,T]} \mathcal{W}_1(\nu_r, \xi_r) \leq C_1 \int_0^T \sup_{r \in [0,s]} \mathcal{W}_1(\nu_r, \xi_r) ds + C_2 \left(I^T(\nu) + \sqrt{I^T(\nu)} \right) + \sqrt{2C_3 I^T(\nu)}.$$

We use Gronwall's Lemma to conclude. \square

4 Lower bound on the exit time of a neighborhood of the stationary state

In this section, we assume that initially, two traits u and v are involved. The stochastic process starts in a state $\nu_0^K = \nu_0^{K,u} + \nu_0^{K,v}$ such that $\nu_0^{K,u}$ is close to $\bar{\xi}^u$ and there exist only a few individuals with trait v . Since the considered initial state is close to the equilibrium $(\bar{\xi}^u, 0)$ and according to Theorem 2.1, the dynamics of the stochastic process ν^K is close to the equilibrium $(\bar{\xi}^u, 0)$ on a finite interval time when K is large. Our aim is to control the exit time of the stochastic process $\nu_t^{K,u}$ from a neighborhood of the stationary solution ξ^u in $M_F(\mathcal{X})$ when K is large and q_K is small. We define the exit time by :

$$\text{for all } \gamma > 0, R_\gamma^K = \inf\{t \geq 0, \mathcal{W}_1(\nu_t^{K,u}, \bar{\xi}^u) \geq \gamma\}. \quad (4.1)$$

Theorem 4.1 gives a lower bound on R_γ^K . The lower bound involves the first time when a new mutation occurs and the first time when the v -population size is larger than a threshold:

$$S_1^K = \inf\{t \geq 0, \exists w \notin \{u, v\}, \nu_t^K(\mathcal{X} \times \{w\}) \neq 0\}, \quad (4.2)$$

$$\text{for all } \epsilon > 0, T_\epsilon^K = \inf\{t \geq 0, \langle \nu_t^{K,v}, \mathbb{1} \rangle \geq \epsilon\}. \quad (4.3)$$

Theorem 4.1. *Suppose that Assumption 1.1 holds and that $H^u > 0$. Let $\gamma > 0$ such that $\gamma < H^u(\kappa^{uu})^{-1}$, and if $H^u\kappa^{vv} - H^v\kappa^{uv} > 0$, γ satisfies also the assumption $\gamma < \left| \frac{H^u}{\kappa^{uu}} - \frac{H^u\kappa^{vv} - H^v\kappa^{uv}}{\kappa^{uu}\kappa^{vv} - \kappa^{vu}\kappa^{uv}} \right|$. Then, there exist $\gamma' > 0$, $\epsilon > 0$, and $V > 0$ such that, if $\nu_0^K = \nu_0^{K,u} + \nu_0^{K,v}$ with $\mathcal{W}_1(\nu_0^{K,u}, \bar{\xi}^u) < \gamma'$ and $\langle \nu_0^{K,v}, \mathbb{1} \rangle < \epsilon$, then*

$$\lim_{K \rightarrow +\infty} \mathbb{P}_{\nu_0^K}(R_\gamma^K > e^{KV} \wedge T_\epsilon^K \wedge S_1^K) = 1.$$

Thus, a well-established monomorphic population u is minimally affected during the emerging of a mutant population v .

The assumptions on the radius γ of the neighborhood ensure that there exists only one steady state in the neighborhood.

The result is proved using ideas similar to the ones of Freidlin and Wentzell [15]. In our framework, the difficulties come from the continuous space motion. Firstly, our processes have values in an infinite dimensional space, thus, the required deterministic results are much more involved, see Subsection 4.1. Secondly, we deal with two kind of randomness : jump process and spatial diffusion process. The end of the section is devoted to the proof of Theorem 4.1.

4.1 Stability for the weak topology

This subsection deals with the deterministic solution to (2.1). We denote by $(\xi_t)_{t \geq 0}$ the solution to equation (2.1) with initial condition $\xi_0 \in M_F(\mathcal{X} \times \{u, v\})$. In this case, $\xi_t \in M_F(\mathcal{X} \times \{u, v\})$ for all $t \geq 0$. We prove that, as long as the size of the v -population density is small, the u -population density stays in a \mathcal{W}_1 -neighborhood of its equilibrium \bar{g}^u .

Proposition 4.2. *Suppose that Assumption 1.1 holds. Let $\gamma > 0$. There exist $\gamma' > 0$ and $\epsilon' > 0$ such that for any $\xi_0 = \xi_0^u \delta_u + \xi_0^v \delta_v$ with $\mathcal{W}_1(\xi_0^u, \bar{\xi}^u) < \gamma'$,*

$$\text{for all } t \leq t_{\epsilon'} = \inf\{t \geq 0, \langle \xi_t^v, \mathbb{1} \rangle > \epsilon'\}, \quad \mathcal{W}_1(\xi_t^u, \bar{\xi}^u) < \gamma/2.$$

The proof of Proposition 4.2 implies two main difficulties. First, using ideas similar to Part 3.3 of [22], we can prove that the solution ξ_t to (2.1) stays close to $\bar{\xi}^u$ if the initial condition admits a density which is close to the density \bar{g}^u of $\bar{\xi}^u$ for the L^2 -distance. However, this is not sufficient since we will deal with discrete measures later. Thus, we need to enlarge the result for \mathcal{W}_1 -distance. Secondly, we are concerned with the trajectories of the u -population process. Even though the v -population size is small, it does have an impact on the death rate of individuals u which we cannot ignore.

The proof is divided into three steps. Firstly, we study how fast a solution with initial condition close to $\bar{\xi}^u$ moves away from $\bar{\xi}^u$ in \mathcal{W}_1 -distance during a small time interval $[0, t_0]$. Then, as $t_0 > 0$, $\xi_{t_0}^u$ admits a density and so, we can compare the \mathcal{W}_1 -distance and the L^2 -distance of the densities between $\xi_{t_0}^u$ and $\bar{\xi}^u$. We finally prove a L^2 -stability result similar to the one of Part 3.3 in [22] but including the v -population process with a small size.

Proof of Proposition 4.2. First we may assume that $\epsilon' \leq 1$ and $\gamma' < \gamma$. Hence, there exists $M > 0$ such that any considered initial state satisfies $\langle \xi_0, \mathbb{1} \rangle < M$.

We fix $t_0 > 0$ and we start with the first step. On the one hand, we can find an upper bound to $\sup_{r \in [0, t]} \langle \xi_r^u, \mathbb{1} \rangle$. Indeed

$$\langle \xi_t^u, \mathbb{1} \rangle \leq \langle \xi_0^u, \mathbb{1} \rangle + \bar{b} \int_0^t \langle \xi_s^u, \mathbb{1} \rangle ds, \quad (4.4)$$

and using Gronwall's Lemma, we deduce that $\sup_{r \in [0, t]} \langle \xi_r^u, \mathbb{1} \rangle \leq M e^{\bar{b}t}$, for all $t \geq 0$. On the other hand, using (2.1) with $\mathcal{U} = \{u, v\}$ and a mild formulation similar to Lemma 4.5 in [5], we find that ξ_t^u satisfies: for any $f \in C^{Lip}(\mathcal{X})$,

$$\begin{aligned} \langle \xi_t^u - \bar{\xi}^u, f \rangle &= \langle \xi_0^u - \bar{\xi}^u, P_t^u f \rangle + \int_0^t \langle \xi_s^u - \bar{\xi}^u, (b^u - d^u - c^{uu} \cdot \bar{\xi}^u) P_{t-s}^u f \rangle ds \\ &\quad + \int_0^t c^{uu} \cdot (\bar{\xi}^u - \xi_s^u) \langle \xi_s^u, P_{t-s}^u f \rangle ds - \int_0^t (c^{uv} \cdot \xi_s^v) \langle \xi_s^u, P_{t-s}^u f \rangle ds. \end{aligned}$$

For any g Lipschitz-continuous, we denote by $\|g\|_{Lip}$ the smallest constant such that $g/\|g\|_{Lip} \in C^{Lip}(\mathcal{X})$. Since $\sup_{t \in [0, t_{\epsilon'}]} \langle \xi_t^v, \mathbb{1} \rangle \leq \epsilon'$, using Lemma 3.5 and the definition of distance \mathcal{W}_1 , we obtain that, for all $t \leq t_{\epsilon'}$,

$$\begin{aligned} |\langle \xi_t^u - \bar{\xi}^u, f \rangle| &\leq \mathcal{W}_1(\xi_0^u, \bar{\xi}^u) + (\bar{b} + \bar{d} + \bar{c} \langle \bar{\xi}^u, \mathbb{1} \rangle) \int_0^t \mathcal{W}_1(\xi_s, \bar{\xi}) ds \\ &\quad + \|c\|_{Lip} \sup_{r \in [0, t]} \langle \xi_r^u, \mathbb{1} \rangle \left(\int_0^t \mathcal{W}_1(\xi_s^u, \bar{\xi}^u) ds + \epsilon' \right). \end{aligned} \quad (4.5)$$

Finally, (4.4), (4.5) and Gronwall's Lemma imply that there exist C_1, C_2 independent of ϵ' and γ' such that

$$\sup_{r \in [0, t_0 \wedge t_{\epsilon'}]} \mathcal{W}_1(\xi_r^u, \bar{\xi}^u) \leq (\mathcal{W}_1(\xi_0^u, \bar{\xi}^u) + \epsilon' C_2) e^{C_1 t_0 \wedge t_{\epsilon'}} \leq (\gamma' + \epsilon' C_2) e^{C_1 t_0}. \quad (4.6)$$

According to (4.6), we have to choose γ' and ϵ' such that $(\gamma' + \epsilon' C_2) e^{C_1 t_0} < \gamma/2$. Note that if for all $\xi_0^v \in M_f(\mathcal{X})$, $t_{\epsilon'} \leq t_0$, the proof of Proposition 4.2 is complete. In what follows, let us assume that $t_{\epsilon'} > t_0$ for the considered initial state ξ_0^v .

The next step is to compare the L^2 -distance and the \mathcal{W}_1 -distance between $\xi_{t_0}^u$ and $\bar{\xi}^u$. According to Theorem 2.1, for any $t_0 > 0$, $\xi_{t_0}^u$ has a Lipschitz-continuous density with respect to Lebesgue measure on \mathcal{X} that we denote by $g_{t_0}^u(x)$. In addition with the fact that $\bar{g}^u \in C^1(\mathcal{X})$, we have

$$\|g_{t_0}^u - \bar{g}^u\|_{L^2}^2 = \int_{\mathcal{X}} (g_{t_0}^u(x) - \bar{g}^u(x))^2 dx \leq \mathcal{W}_1(\xi_{t_0}^u, \bar{\xi}^u) (\|g_{t_0}^u\|_{Lip} + \|\bar{g}^u\|_{Lip}). \quad (4.7)$$

Let us bound $\|g_{t_0}^u\|_{Lip}$ from above. For any $t > 0$, we define $h_t^u(x) = g_t^u(x) \exp(\int_0^t (c^{uu} \cdot g_s^u + c^{uv} \cdot \xi_s^v) ds)$. The exponent of the exponential term is positive and independent of x , thus $\|g_{t_0}^u\|_{Lip} \leq$

$\|h_{t_0}^u\|_{Lip}$. Furthermore, according to Part 4 of chapter 5 in [16], $h_{t_0}^u(x) = \int_{\mathcal{X}} \Gamma_{t_0}(x, y) \xi_0(dy)$ where Γ is the fundamental solution to the system

$$\begin{cases} \partial_t \Gamma = m^u \Delta \Gamma + (b^u(x) - d^u(x)) \Gamma \text{ on } \mathcal{X} \times \mathbb{R}^+, \\ \partial_n \Gamma = 0 \text{ on } \partial \mathcal{X} \times \mathbb{R}^+, \\ \Gamma(0, dx) = \xi_0(dx). \end{cases}$$

As $t_0 > 0$, $\|\Gamma_{t_0}\|_{L^\infty(\mathcal{X})}$ and $\|\nabla \Gamma_{t_0}\|_{L^\infty(\mathcal{X})}$ are bounded from above and there exists C_3 such that

$$\|g_{t_0}^u\|_{Lip} \leq \|h_{t_0}^u\|_{Lip} \leq C_3 \langle \xi_0^u, \mathbf{1} \rangle \leq C_3 M, \quad (4.8)$$

where M has been defined in the beginning of the proof. Combining (4.6), (4.7), (4.8) and the fact that $\mathcal{W}_1(\xi_0^u, \bar{\xi}^u) < \gamma'$, we find $C_4(\epsilon', \gamma') > 0$ such that

$$\begin{aligned} \|g_{t_0}^u - \bar{g}^u\|_{L^2} &\leq ((\mathcal{W}_1(\xi_0^u, \bar{\xi}^u) + C_2 \epsilon') e^{C_1 t_0} (C_3 M + \|\bar{g}^u\|_{Lip}))^{1/2} \\ &\leq ((\gamma' + C_2 \epsilon') e^{C_1 t_0} (C_3 M + \|\bar{g}^u\|_{Lip}))^{1/2} := C_4(\epsilon', \gamma'). \end{aligned}$$

We now deal with the last step of the proof. We define $(\lambda_k, A_k)_{k \geq 1}$ the spectral basis for the operator $m^u \Delta_x + (b^u - d^u)$, with Neumann boundary condition such that $(\lambda_k)_{k \geq 1}$ is a non-decreasing sequence with $H^u := \lambda_1 > \lambda_2 \geq \lambda_3 \dots$ and $(A_k)_{k \geq 1}$ is an orthonormal basis of $L^2(\mathcal{X})$. Note that \bar{g}^u is equal to $\|\bar{g}^u\|_{L^2} A_1$. Let us express $g^u - \bar{g}^u$ in the basis $(A_k)_{k \geq 1}$

$$g_t^u(x) = \bar{g}^u(x) + \sum_{i=1}^{+\infty} \alpha_i(t) A_i(x),$$

and denote for all $t \in \mathbb{R}^+$,

$$\beta(t) := 1 + \frac{\alpha_1(t)}{\|\bar{g}^u\|_{L^2}}.$$

From (2.1) and the representation of $g^u - \bar{g}^u$ and $\partial_t g^u$ with respect to the basis $(A_k)_{k \geq 1}$, we find the following dynamical system

$$\partial_t \alpha_k(t) = \alpha_k(t) \left(\lambda_k - H^u - \int_{\mathcal{X}} c^{uv} g_t^v - \int_{\mathcal{X}} c^{uu} (g_t^u - \bar{g}^u) \right), \text{ for all } k \geq 2, \quad (4.9)$$

$$\partial_t \beta(t) = \beta(t) \left(H^u - \int_{\mathcal{X}} c^{uv} g_t^v - \sum_{i=2}^{+\infty} \alpha_i(t) \left(\int_{\mathcal{X}} c^{uu} A_i \right) - H^u \beta(t) \right). \quad (4.10)$$

The last step of the proof consists in proving that if ϵ' and γ' are sufficiently small such that

$$\max \left\{ 1, \frac{\|\bar{g}^u\|_{L^2}}{H^u} \right\} \cdot (\epsilon' \bar{c} + 3 \|c^{uu}\|_{L^2} C_4(\epsilon', \gamma')) < \min \left\{ \frac{H^u - \lambda_2}{2}, \frac{\gamma}{2 \|\mathbf{1}\|_{L^2}} \right\}, \quad (4.11)$$

then for all $t \leq t_{\epsilon'}$, $\mathcal{W}_1(g_t^u, \bar{g}^u) < \gamma/2$. Let us fix $K > 0$ such that

$$\epsilon' \bar{c} + 3 \|c^{uu}\|_{L^2} C_4(\epsilon', \gamma') \leq K/2 < (H^u - \lambda_2)/2, \quad (4.12)$$

and

$$t_{max} = \inf\{t \geq t_0, |\int_{\mathcal{X}} c^{uu}(g_t^u - \bar{g}^u)| \geq K\}.$$

Let us prove that $t_{max} \geq t_{\epsilon'}$. At $t = t_0$, $|\int_{\mathcal{X}} c^{uu}(g_{t_0}^u - \bar{g}^u)| \leq \|c^{uu}\|_{L^2} C_4(\epsilon', \gamma') \leq K/2 < K$. A continuity argument implies that $t_{max} > t_0$. Then, using (4.9), we deduce that for any $k \geq 2$, for any $t_0 \leq t \leq t_{max} \wedge t_{\epsilon'}$,

$$\alpha_k(t)^2 \leq \alpha(t_0)^2 \exp(2(\lambda_k - H^u + K)(t - t_0)) \leq \alpha(t_0)^2.$$

Thus

$$\begin{aligned} \left| \sum_{i=2}^{+\infty} \alpha_i(t) \int_{\mathcal{X}} (c^{uu} A_i) \right| &\leq \left(\sum_{i=2}^{+\infty} \alpha_i(t)^2 \right)^{1/2} \left(\sum_{i=2}^{+\infty} \left(\int_{\mathcal{X}} c^{uu} A_i \right)^2 \right)^{1/2} \\ &\leq \left(\sum_{i=1}^{+\infty} \alpha_i(t_0)^2 \right)^{1/2} \left(\sum_{i=1}^{+\infty} \left(\int_{\mathcal{X}} c^{uu} A_i \right)^2 \right)^{1/2} \leq C_4(\epsilon', \gamma') \|c^{uu}\|_{L^2}. \end{aligned} \quad (4.13)$$

Inserting (4.13) in Equation (4.10) implies that, for all $t \leq t_{max} \wedge t_{\epsilon'}$,

$$\begin{aligned} \beta(t) (H^u - \epsilon' \bar{c} - C_4(\epsilon', \gamma') \|c^{uu}\|_{L^2} - H^u \beta(t)) &\leq \partial_t \beta(t) \\ &\leq \beta(t) (H^u + \epsilon' \bar{c} + C_4(\epsilon', \gamma') \|c^{uu}\|_{L^2} - H^u \beta(t)). \end{aligned}$$

Moreover, $|\beta(t_0) - 1| \|\bar{g}^u\|_{L^2} \leq C_4(\epsilon', \gamma')$. Using properties of logistic equations, we deduce that for all $t_0 \leq t \leq t_{max} \wedge t_{\epsilon'}$,

$$|\alpha_1(t)| = |\beta(t) - 1| \|\bar{g}^u\|_{L^2} \leq \frac{\|\bar{g}^u\|_{L^2}}{H^u} (\epsilon' \bar{c} + C_4(\epsilon', \gamma') \|c^{uu}\|_{L^2}) + C_4(\epsilon', \gamma'). \quad (4.14)$$

Thus, for all $t_0 \leq t \leq t_{max} \wedge t_{\epsilon'}$, (4.13) and (4.14) imply that

$$\begin{aligned} \left| \int_{\mathcal{X}} c^{uu}(g_t^u - \bar{g}^u) \right| &\leq |\alpha_1(t)| \frac{H^u}{\|\bar{g}^u\|_{L^2}} + \left| \sum_{i=2}^{+\infty} \alpha_i(t) \int_{\mathcal{X}} (c^{uu} A_i) \right| \\ &\leq \epsilon' \bar{c} + 2C_4(\epsilon', \gamma') \|c^{uu}\|_{L^2} + \frac{H^u}{\|\bar{g}^u\|_{L^2}} C_4(\epsilon', \gamma'). \end{aligned}$$

By definition, $H^u = \int_{\mathcal{X}} c^{uu} \bar{g}^u \leq \|c^{uu}\|_{L^2} \|\bar{g}^u\|_{L^2}$ and (4.12) gives for all $t \leq t_{max} \wedge t_{\epsilon'}$,

$$\left| \int_{\mathcal{X}} c^{uu}(g_t^u - \bar{g}^u) \right| \leq \epsilon' \bar{c} + 3C_4(\epsilon', \gamma') \|c^{uu}\|_{L^2} \leq \frac{K}{2}. \quad (4.15)$$

Inequality (4.15) implies that $t_{max} \wedge t_{\epsilon'} < t_{max}$, that is, $t_{max} \wedge t_{\epsilon'} = t_{\epsilon'}$ and all inequalities proved above are true for all $t \leq t_{\epsilon'}$. In addition, (4.11), (4.13) and (4.14) imply that, for all $t \leq t_{\epsilon'}$,

$$\begin{aligned} \mathcal{W}_1(g_t^u, \bar{g}^u) &\leq \|g_t^u - \bar{g}^u\|_{L^2} \|\mathbf{1}\|_{L^2} \leq \|\mathbf{1}\|_{L^2} \left(|\alpha_1(t)| + \left| \sum_{i=1}^{+\infty} \alpha_i(t)^2 \right|^{1/2} \right) \\ &\leq \frac{\|\mathbf{1}\|_{L^2} \|\bar{g}^u\|_{L^2}}{H^u} (\epsilon' \bar{c} + 3C_4(\epsilon', \gamma') \|c^{uu}\|_{L^2}) \leq \frac{\gamma}{2}. \end{aligned}$$

That ends the proof. \square

4.2 Exit time

This subsection is devoted to the proof of Theorem 4.1. We split the proof into three lemmas similar to the ones in Dembo and Zeitouni [9].

Let $\gamma > 0$ satisfying the assumptions of Theorem 4.1. We consider ϵ' and γ' as in Proposition 4.2 and set $\epsilon = \frac{\epsilon'}{2}$ and $\rho = \frac{\gamma'}{3} < \gamma$. R_γ^K , S_1 and T_ϵ^K have been defined by (4.1), (4.2) and (4.3) and let us define

$$\tau = \inf\{t \geq 0, \mathcal{W}_1(\nu_t^{K,u}, \bar{\xi}^u) \notin [\rho, \gamma]\}.$$

Lemma 4.3. *Under Assumption 1.1, we have*

$$\lim_{t \rightarrow +\infty} \limsup_{K \rightarrow +\infty} \frac{1}{K} \log \sup_{\nu_0^K \in (B(\bar{\xi}^u, \gamma) \times B(0, \epsilon)) \cap M_F^K} \mathbb{P}_{\nu_0^K}(\tau \wedge T_\epsilon^K \wedge S_1^K > t) = -\infty.$$

That is, the probability that the process $\nu^{K,u}$ stays a long time in the ring $B(\bar{\xi}^u, \gamma) \setminus B(\bar{\xi}^u, \rho)$ is exponentially small. The proof requires a comparison with the deterministic paths of Equation (2.1), the difficulty is to prove that there exists a finite time after which all deterministic paths starting in the ring are out of the ring. The fact that the probability is exponentially small is a consequence of Theorem 3.1.

Lemma 4.4. *Under Assumption 1.1, there exists $V > 0$ such that*

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \sup_{\nu_0^K \in (B(\bar{\xi}^u, \gamma') \times B(0, \epsilon)) \cap M_F^K} \mathbb{P}_{\nu_0^K}(\mathcal{W}_1(\nu_\tau^{K,u}, \bar{\xi}^u) \geq \gamma, \tau \leq T_\epsilon^K \wedge S_1^K) \leq -V.$$

Once again, the proof is based on a comparison with the deterministic paths, it is a consequence of Proposition 4.2 and Theorem 3.1.

Lemma 4.5. *Under Assumption 1.1, for all $C > 0$, there exists $T(C, \rho) > 0$ such that,*

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \sup_{\nu_0^{K,u} \in B(\bar{\xi}^u, \gamma) \cap M_F^K} \mathbb{P}_{\nu_0^K} \left(\sup_{t \in [0, T(C, \rho)]} \mathcal{W}_1(\nu_t^{K,u}, \nu_0^{K,u}) \geq \rho \right) < -C.$$

This lemma means that there exists a small time interval during which any process stays close from its starting point with an exponentially high probability. The stochastic process includes jump parts and diffusive parts. Thus, we study not only the size of the population process during a small time interval, but also a sum of reflected diffusion processes.

Theorem 4.1 is proved using the two last lemmas 4.4 and 4.5. We do not detail its proof as it can be adapted from [9]. The main change is that the proof has to be done on $\{t \leq T_\epsilon^K \wedge S_1^K\}$ to ensure that the v -process size $\langle \nu^{K,v}, \mathbb{1} \rangle$ is small and that no other mutation appears, but Lemma 4.4 is sufficient to circumvent this difficulty.

There remains to prove the three lemmas.

Proof of Lemma 4.3. Remark that if $\mathcal{W}_1(\nu_0^{K,u}, \bar{\xi}^u) < \rho$ for all $K \geq 1$, the result is obvious. Otherwise, let us define the following set:

$$\mathcal{A}(T) = \text{adh}\{\nu \in \mathbb{D}([0, T], M_F(\mathcal{X} \times \{u, v\})), \forall t \in [0, T], \mathcal{W}_1(\nu_t^u, \bar{\xi}^u) \in]\rho, \gamma[\text{ and } \langle \nu_t^v, \mathbb{1} \rangle < \epsilon\}, \quad (4.16)$$

where $\text{adh}S$ is the closure of the set S . Remark that $\{\nu^K \in \mathcal{A}(T)\} = \{\tau \wedge T_\epsilon^K \wedge S_1^K > T\}$ a.s. and that the set $\mathcal{C} = \text{adh}(B(\bar{\xi}^u, \gamma) \times B(0, \epsilon))$ is a compact set of $M_F(\mathcal{X} \times \{u, v\})$ as \mathcal{X} is bounded. By applying Theorem 3.1 with the closed set $\mathcal{A}(T)$ and the initial compact set \mathcal{C} , we find

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \sup_{\nu_0^K \in \mathcal{C} \cap M_F^K} \mathbb{P}_{\nu_0^K}(\nu \in \mathcal{A}(T)) \leq - \inf_{\nu \in \mathcal{A}(T)} I^T(\nu).$$

Thus, the lemma will be proved if we show that $\inf_{\nu \in \mathcal{A}(T)} I^T(\nu) \rightarrow +\infty$ as T tends to $+\infty$. To this aim, we will first show that any solution to (2.1) cannot belong to $\mathcal{A}(T_0)$ if T_0 is large enough. Precisely, we set $\delta \in]0, \rho/2[$, and we will prove that there exists $T_0 > 0$, such that any $(\xi_t)_{t \geq 0}$ solution to (2.1) with an initial condition satisfying $\mathcal{W}_1(\xi_0^u, \bar{\xi}^u) \in]\rho, \gamma[$ and $\langle \xi_0^v, \mathbb{1} \rangle < \epsilon$ satisfies

$$\inf_{\nu \in \mathcal{A}(T_0)} \sup_{t \in [0, T_0]} \mathcal{W}_1(\xi_t, \nu_t) \geq \delta. \quad (4.17)$$

Assume that (4.17) holds. Since $\sup_{\nu \in \mathcal{A}(T_0)} \sup_{t \in [0, T_0]} \langle \nu, \mathbb{1} \rangle$ is bounded, we can use Proposition 3.4 to deduce that there exists $C > 0$ such that for any $\nu \in \mathcal{A}(T_0)$, $\delta \leq C(I^{T_0}(\nu) + \sqrt{I^{T_0}(\nu)})$. As $x \mapsto x + \sqrt{x}$ is a bijective function from \mathbb{R}^+ to \mathbb{R}^+ , we find a constant $C(\delta) > 0$ which is a lower bound on $I^{T_0}(\nu)$ and $\inf_{\nu \in \mathcal{A}(T_0)} I^{T_0}(\nu) \geq C(\delta)$. Finally, for $T > T_0$ and $\nu \in \mathcal{A}(T)$, we decompose ν as a sum of $n := \lceil T/T_0 \rceil$ terms : using Chasles' Theorem and the non-variational formulation (3.8) of $I^T > 0$, we find a sequence $(\nu^i)_{i=1..n} \in \mathcal{A}(T_0)$ such that

$$I^T(\nu) \geq \sum_{i=1}^n I^{T_0}(\nu^i) \geq C(\delta)n \xrightarrow{T \rightarrow +\infty} +\infty.$$

There remains to prove that (4.17) holds. Let $\mathcal{W}_1(\xi_0^u, \bar{\xi}^u) \in]\rho, \gamma[$ and $\langle \xi_0^v, \mathbb{1} \rangle < \epsilon$, and let $(\xi_t)_{t \geq 0}$ be the solution to (2.1) with initial condition $\xi_0 = \xi_0^u \delta_u + \xi_0^v \delta_v$. Using Theorems 1.2 and 1.4 in [22], we know that $\xi_t \in M_F(\mathcal{X} \times \{u, v\})$ converges to a stationary state. Theorem 1.2 in [22] implies that this stationary state is either $\bar{\xi}^u \delta_u$ or $\bar{\xi}^v \delta_v$ or a state with coexistence but the assumptions on γ insure that those stationary states do not belong to $(B(\bar{\xi}^u, \gamma) \setminus B(\bar{\xi}^u, \rho)) \times B(0, \epsilon)$. Therefore, there exists T_{ξ_0} such that

$$\mathcal{W}_1(\xi_{T_{\xi_0}}^u, \bar{\xi}^u) \notin [\gamma + 2\delta, \rho - 2\delta] \text{ or } \langle \xi_{T_{\xi_0}}^v, \mathbb{1} \rangle \geq \epsilon + 2\delta. \quad (4.18)$$

However T_{ξ_0} depends on ξ_0 . Thus, we will use a compactness argument to find a uniform time and conclude. First, using Gronwall's Lemma, we obtain : for $M \geq \langle \bar{\xi}^u, \mathbb{1} \rangle + 2\gamma + 2\epsilon$,

$$\sup_{t \in [0, T_{\xi_0}]} \langle \xi_t, \mathbb{1} \rangle \leq \langle \xi_0, \mathbb{1} \rangle e^{\bar{b}T_{\xi_0}} \leq M e^{\bar{b}T_{\xi_0}}. \quad (4.19)$$

Then we show that two solutions $(\xi_t)_{t \geq 0}$ and $(\zeta_t)_{t \geq 0}$ to (2.1) which are close initially stay close during a short time. Indeed, using the mild equation of Lemma 4.5 in [5], for any $f \in C^{Lip}(\mathcal{X} \times \{u, v\})$

$$\langle \xi_t - \zeta_t, f \rangle = \langle \xi_0 - \zeta_0, P_t f \rangle + \int_0^t \langle \xi_s - \zeta_s, (b - d - c \cdot \xi_s) P_{t-s} f \rangle ds - \int_0^t \langle \xi_s, c \cdot (\xi_s - \zeta_s) P_{t-s} f \rangle ds.$$

According to Lemma 3.5 and Assumptions 1.1, we can find a constant C_1 such that $(b - d - c \cdot \xi_t) P_{t-s} f / C_1$ and c / C_1 belong to $C^{Lip}(\mathcal{X} \times \{u, v\})$. We deduce that for all $t \leq T_{\xi_0}$,

$$|\langle \xi_t - \zeta_t, \varphi \rangle| \leq \mathcal{W}_1(\xi_0, \zeta_0) + C_1 \int_0^t \sup_{r \in [0, s]} \mathcal{W}_1(\xi_r, \zeta_r) ds + C_1 \sup_{r \in [0, T_{\xi_0}]} \langle \xi_r, \mathbb{1} \rangle \int_0^t \sup_{r \in [0, s]} \mathcal{W}_1(\xi_r, \zeta_r) ds. \quad (4.20)$$

Using Gronwall's Lemma and (4.19), we conclude that there exists a constant $C(T_{\xi_0})$ such that

$$\sup_{r \in [0, T_{\xi_0}]} \mathcal{W}_1(\xi_r, \zeta_r) \leq C(T_{\xi_0}) \mathcal{W}_1(\xi_0, \zeta_0).$$

Choosing $\alpha_{\xi_0} = \delta / C(T_{\xi_0})$, we find for all ζ_0 with $\mathcal{W}_1(\xi_0, \zeta_0) < \alpha_{\xi_0}$, $\sup_{r \in [0, T_{\xi_0}]} \mathcal{W}_1(\xi_r, \zeta_r) < \delta$. In addition with (4.18) and (4.16), we find that for all ζ_0 with $\mathcal{W}_1(\zeta_0, \xi_0) < \alpha_{\xi_0}$,

$$\sup_{t \in [0, T_{\xi_0}]} \sup_{\nu \in \mathcal{A}(T_0)} \mathcal{W}_1(\zeta_t, \nu_t) \geq \delta.$$

Remark that $\text{adh}((B(\bar{\xi}^u, \gamma) \setminus B(\bar{\xi}^u, \rho)) \times B(0, \epsilon))$ is a compact set of $M_F(\mathcal{X} \times \{u, v\})$ as \mathcal{X} is bounded. It is covered by $\cup_{\xi_0 \in (B(\bar{\xi}^u, \gamma) \setminus B(\bar{\xi}^u, \rho)) \times B(0, \epsilon)} B(\xi_0, \alpha_{\xi_0})$. We extract a finite cover $\cup_{i=1}^n B(\xi_0^i, \alpha_{\xi_0^i})$. Finally, defining $T_0 = \max_{i=1..n} T_{\xi_0^i}$, we conclude : for any ξ_0 with $\mathcal{W}_1(\xi_0^u, \bar{\xi}^u) \in]\rho, \gamma[$ and $\langle \xi_0^v, \mathbb{1} \rangle < \epsilon$, we have

$$\inf_{\nu \in \mathcal{A}(T_0)} \sup_{t \in [0, T_0]} \mathcal{W}_1(\xi_t, \nu_t) \geq \delta.$$

Proof of Lemma 4.3 is now complete. \square

Proof of Lemma 4.4. Lemma 4.3 gives T_1 such that

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \sup_{\nu_0^K \in (B(\bar{\xi}^u, \gamma) \times B(0, \epsilon)) \cap M_F^K} \mathbb{P}_{\nu_0^K}(\tau \wedge T_\epsilon^K \wedge S_1^K > T_1) \leq -1. \quad (4.21)$$

Thus we limit our study to the time interval $[0, T_1]$. Since the considered initial states satisfy $\langle \nu_0^K, \mathbb{1} \rangle \leq (\langle \bar{\xi}^u, \mathbb{1} \rangle + 2\gamma + 2\epsilon)$, using Lemma 3.2, we find $N > 0$ such that

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \sup_{\nu_0^K \in (B(\bar{\xi}^u, \gamma) \times B(0, \epsilon)) \cap M_F^K} \mathbb{P}_{\nu_0^K} \left(\sup_{t \in [0, T_1]} \langle \nu_t^K, \mathbb{1} \rangle \geq N \right) \leq -1. \quad (4.22)$$

Let $M \geq (\langle \bar{\xi}^u, \mathbf{1} \rangle + 2\gamma + 2\epsilon) \vee N$ and

$$\mathcal{A} = \left\{ \nu \in \mathbb{D}([0, T_1], M_F(\mathcal{X} \times \{u, v\})) \mid \exists t \in [0, T_1], \mathcal{W}_1(\nu_t^u, \bar{\xi}^u) < \gamma, \sup_{t \in [0, T_1]} \langle \nu_t^v, \mathbf{1} \rangle < \epsilon \right. \\ \left. \text{and } \sup_{t \in [0, T_1]} \langle \nu_t, \mathbf{1} \rangle < M \right\}. \quad (4.23)$$

For all ν_0^K , and K ,

$$\begin{aligned} \mathbb{P}_{\nu_0^K}(\mathcal{W}_1(\nu_\tau^{K,u}, \bar{\xi}^u) \geq \gamma, \tau \leq T_\epsilon^K \wedge S_1^K) \\ \leq \mathbb{P}_{\nu_0^K}(\tau \leq T_1, \sup_{t \in [0, T_1]} \langle \nu_t^K, \mathbf{1} \rangle \geq M) + \mathbb{P}_{\nu_0^K}(\tau \leq T_1, \nu^K \in \mathcal{A}) + \mathbb{P}_{\nu_0^K}(\tau > T_1, \tau \leq T_\epsilon^K \wedge S_1^K) \\ \leq \mathbb{P}_{\nu_0^K}(\sup_{t \in [0, T_1]} \langle \nu_t^K, \mathbf{1} \rangle \geq M) + \mathbb{P}_{\nu_0^K}(\nu^K \in \mathcal{A}) + \mathbb{P}_{\nu_0^K}(\tau \wedge T_\epsilon^K \wedge S_1^K > T_1). \end{aligned}$$

Then, we use Theorem 3.1, the definition of \mathcal{A} (4.23), (4.21), and (4.22) to find

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \sup_{\nu_0^K \in \mathcal{C} \cap M_F^K} \mathbb{P}_{\nu_0^K}(\nu_\tau^K \in B^c(\bar{\xi}^u, \gamma)) \leq \max \left\{ -1, - \inf_{\nu \in \mathcal{A}, \nu_0 \in \mathcal{C}} I^{T_1}(\nu) \right\},$$

where $\mathcal{C} = \text{adh}(B(\bar{\xi}^u, 3\rho) \times B(0, \epsilon))$. There remains to prove that the r.h.s. is strictly negative. Proposition 4.2 implies that any solution $(\xi_t)_{t \geq 0}$ to (2.1) with $\xi_0 \in \mathcal{C}$ satisfies the property :

$$\text{if } \sup_{t \in [0, T_1]} \langle \xi_t^v, \mathbf{1} \rangle < \epsilon' = 2\epsilon, \text{ then } \mathcal{W}_1(\xi_t^u, \bar{\xi}^u) < \gamma/2.$$

We deduce immediately that for any $\nu \in \mathcal{A}$, with $\nu_0 \in \mathcal{C}$, if ξ is the solution to (2.1) with $\xi_0 := \nu_0$,

$$\sup_{t \in [0, T_1]} \mathcal{W}_1(\nu_t, \xi_t) \geq \sup_{t \in [0, T_1]} \max\{\mathcal{W}_1(\nu_t^u, \xi_t^u), \mathcal{W}_1(\nu_t^v, \xi_t^v)\} \geq \frac{\gamma}{2} \wedge \epsilon.$$

We conclude the proof using Proposition 3.4 and a compactness argument similar to the one at the end of the previous proof. \square

Proof of Lemma 4.5. Let us fix $f \in C^{Lip}(\mathcal{X})$ and study the following difference, using the construction of the process $\nu^{K,u}$:

$$|\langle \nu_t^{K,u}, f \rangle - \langle \nu_0^{K,u}, f \rangle| \leq \frac{1}{K} \left(\sum_{i \in N_t^{\text{notdead}}} |f(X_t^i) - f(X_0^i)| + \sum_{i \in N_t^{\text{dead}}} |f(X_0^i)| + \sum_{i \in N_t^{\text{born}}} |f(X_t^i)| \right),$$

where N_t^{notdead} is the set of indices of the individuals with trait u alive at time 0 and not dead during $[0, t]$; N_t^{dead} represents the set of indices of the individuals with trait u alive at time 0 and dead during $[0, t]$; and N_t^{born} is the set of indices of the individuals with trait u born during $]0, t]$. As $f \in C^{Lip}(\mathcal{X})$,

$$|\langle \nu_t^{K,u}, f \rangle - \langle \nu_0^{K,u}, f \rangle| \leq \frac{1}{K} \left(\sum_{i \in N_t^{\text{notdead}}} |X_t^i - X_0^i| \right) + \frac{N_t^{\text{dead}} + N_t^{\text{born}}}{K}.$$

As this is true for all $f \in C^{Lip}(\mathcal{X})$, we find the same upper bound for $\mathcal{W}_1(\nu_t^{K,u}, \nu_0^{K,u})$. We use now the stopping time

$$\tau_N^K = \inf\{t > 0, |\langle \nu_t^{K,u}, \mathbf{1} \rangle| > N\}.$$

On $\{\tau_N^K \geq t\}$, $\mathcal{W}_1(\nu_t^{K,u}, \nu_0^{K,u})$ is stochastically bounded by $\frac{1}{K} \sum_{i=1}^{KN} |X_t^i - X_0^i| + \frac{\mathcal{P}(t)}{K}$, where $\{(X_t^i)_{t \geq 0}\}_{i \in \{1..KN\}}$ are KN independent reflected diffusion processes driven by Equation (1.2) with the diffusion coefficient m^u and $(\mathcal{P}(t))_{t \geq 0}$ is a Poisson process with intensity $(\bar{b} + \bar{d} + N\bar{c})KN$. Finally,

$$\begin{aligned} & \mathbb{P}\left(\sup_{t \in [0, T]} \mathcal{W}_1(\nu_t^{K,u}, \nu_0^{K,u}) \geq \rho\right) \\ & \leq \mathbb{P}(\tau_N^K \leq T) + \mathbb{P}\left(\tau_N^K \geq T, \sup_{t \in [0, T]} \mathcal{W}_1(\nu_t^{K,u}, \nu_0^{K,u}) \geq \rho\right), \\ & \leq \mathbb{P}(\tau_N^K \leq T) + \mathbb{P}\left(\frac{1}{K} \sum_{i=1}^{KN} \sup_{t \in [0, T]} |X_t^i - X_0^i| \geq \frac{\rho}{2}\right) + \mathbb{P}\left(\sup_{t \in [0, T]} \frac{\mathcal{P}(t)}{K} \geq \frac{\rho}{2}\right), \end{aligned} \quad (4.24)$$

Using Lemma 3.2, we can fix $N \in \mathbb{N}$ such that

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \mathbb{P}(\tau_N^K \leq T) \leq -C.$$

Let us now consider the second term of (4.24). For $\bar{x} = (x^1, \dots, x^{KN}) \in \mathcal{X}^{KN}$, we denote the probability under which (X_0^1, \dots, X_0^{KN}) is equal to \bar{x} by $\mathbb{P}_{\bar{x}}$. Let Υ be the stopping time $\Upsilon = \inf\{s \geq 0, \sum_{i=1}^{KN} |X_s^i - X_0^i| \geq \rho K/2\}$. Using the Markov property, we find

$$\begin{aligned} & \mathbb{P}_{\bar{x}}\left(\sup_{t \in [0, T]} \sum_{i=1}^{KN} |X_t^i - X_0^i| \geq \frac{\rho K}{2}\right) = \mathbb{P}_{\bar{x}}(\Upsilon \leq T) \\ & \leq \mathbb{P}_{\bar{x}}\left(\Upsilon \leq T, \sum_{i=1}^{KN} |X_T^i - X_0^i| \geq \frac{\rho K}{4}\right) + \mathbb{P}_{\bar{x}}\left(\Upsilon \leq T, \sum_{i=1}^{KN} |X_T^i - X_{\Upsilon}^i| \geq \frac{\rho K}{4}\right) \\ & \leq \mathbb{P}_{\bar{x}}\left(\sum_{i=1}^{KN} |X_T^i - X_0^i| \geq \frac{\rho K}{4}\right) + \mathbb{E}_{\bar{x}}\left[\int_0^T \mathbb{P}_{\bar{X}_s}\left(\sum_{i=1}^{KN} |X_{T-s}^i - X_0^i| \geq \frac{\rho K}{4}\right) \mathbf{1}_{\Upsilon \in ds}\right] \\ & \leq 2 \sup_{\bar{y} \in \mathcal{X}^{KN}, s \in [0, T]} \mathbb{P}_{\bar{y}}\left(\sum_{i=1}^{KN} |X_s^i - X_0^i| \geq \frac{\rho K}{4}\right). \end{aligned} \quad (4.25)$$

The aim is thus to find an upper bound on the last term for any $\bar{x} \in \mathcal{X}^{KN}$ and any $s \in [0, T]$. Using Markov's inequality,

$$\mathbb{P}_{\bar{x}}\left(\sum_{i=1}^{KN} |X_s^i - X_0^i| \geq \frac{\rho K}{4}\right) \leq e^{-\frac{K\rho}{4\sqrt{T}}} \prod_{i=1}^{KN} \mathbb{E}_{x^i}\left[e^{\frac{|X_s^i - x^i|}{\sqrt{s}}}\right]. \quad (4.26)$$

If we denote the kernel of the semigroup P_s^u of the reflected diffusion process by $p_s^u(x, y)$, Part 3 in [32] and the fact that $\bar{\mathcal{X}}$ is compact imply that there exist two positive constants C_1, C_2 such that for any $x, y \in \mathcal{X}$,

$$p_s^u(x, y) \leq \frac{C_1}{s^{d/2}} e^{-\frac{|x-y|^2}{C_2 s}}. \quad (4.27)$$

Thus, using (4.27) then a change of variables, we find that there exists $C_3 > 0$ independent from s such that

$$\mathbb{E}_{x^i} \left[e^{\frac{|X_s^i - x^i|}{\sqrt{s}}} \right] \leq \int_{\mathbb{R}^d} \frac{C_1}{s^{d/2}} e^{-\frac{|x-y|^2}{C_2 s}} e^{\frac{|y-x|}{\sqrt{s}}} dy \leq \int_{\mathbb{R}^d} C_1 e^{-\frac{|z|^2}{C_2}} e^{|z|} dz = C_3 < +\infty. \quad (4.28)$$

We deduce with the last line in (4.25), (4.26) and (4.28) that

$$\mathbb{P}_{\bar{x}} \left(\sup_{t \in [0, T]} \sum_{i=0}^{KN} |X_t^i - X_0^i| \geq \frac{\rho K}{2} \right) \leq 2e^{-K \left(\frac{\rho}{4\sqrt{T}} - \ln(C_3)N \right)},$$

where $\frac{\rho}{4\sqrt{T}} - \ln(C_3)N$ tends to $+\infty$ when T tends to 0, i.e. there exists T_1 such that for all $T \leq T_1$, $\frac{\rho}{4\sqrt{T}} - \ln(C_3)N \geq C$.

Finally, concerning the third term of (4.24), let $r = \log \left(\frac{\rho}{2(\bar{b} + \bar{d} + \bar{c}N)NT} \right)$,

$$\begin{aligned} \mathbb{P} \left(\sup_{t \in [0, T]} \frac{\mathcal{P}(t)}{K} \geq \frac{\rho}{2} \right) &\leq \mathbb{P} \left(e^{r\mathcal{P}(T)} \geq e^{r\frac{K\rho}{2}} \right) \\ &\leq e^{-\frac{\rho r K}{2}} \mathbb{E}[e^{r\mathcal{P}(T)}] \\ &\leq e^{-K \left(\frac{\rho r}{2} - (\bar{b} + \bar{d} + \bar{c}N)NT(e^r - 1) \right)}, \\ &\leq e^{-K \left(\frac{\rho}{2} \left[\log \left(\frac{\rho}{2(\bar{b} + \bar{d} + \bar{c}N)NT} \right) - 1 \right] + (\bar{b} + \bar{d} + \bar{c}N)NT \right)}. \end{aligned}$$

There exists $T_2 \leq T_1$ such that for all $T \leq T_2$, $\frac{\rho}{2} \left[\log \left(\frac{\rho}{2(\bar{b} + \bar{d} + \bar{c}N)NT} \right) - 1 \right] + (\bar{b} + \bar{d} + \bar{c}N)NT \geq C$. We conclude the proof : for all $T \leq T_2$,

$$\limsup_{K \rightarrow +\infty} \frac{1}{K} \log \mathbb{P} \left(\sup_{t \in [0, T]} \mathcal{W}_1(\nu_T^{K,u}, \nu_0^{K,u}) \geq \rho \right) \leq -C.$$

□

5 Survival probability for a branching diffusion process

In this part, we study a distinct model which is a branching diffusion process. This model is correlated with the other one as explained at the end of Section 2. Any individual is characterized by its location $X_t^i \in \mathcal{X}$ which is described as before by a diffusion process normally reflected at the boundary of \mathcal{X} , (1.2), with the diffusion coefficient $m > 0$. Moreover, each individual with location $x \in \mathcal{X}$ gives birth to a new individual at rate $b(x)$ and dies at rate $d(x)$. Those rates are assumed to have some regularity that we detail.

Assumption 5.1. b, d are two Lipschitz functions, b is a positive function and there exists $\underline{b}, \bar{b}, \bar{d}$ such that for all $x \in \mathcal{X}$, $\underline{b} < b(x) \leq \bar{b}$ and $0 \leq d(x) \leq \bar{d}$. Moreover, d is a non-zero function.

Let M_t denote the number of individuals at time t . We describe the dynamics of the diffusion process at each time by the finite measure

$$\eta_t = \sum_{i=1}^{M_t} \delta_{X_t^i}.$$

The aim of this part is to describe the survival probability of the population using its parameters. Let

$$\Upsilon_0 = \inf\{t \geq 0, M_t = 0\}.$$

The first Theorem concerns the survival probability of the population assuming that, initially, there is only one individual at location x . This probability is characterized as a solution to an elliptic differential equation on the location space \mathcal{X} . Remark that the location of the first individual plays a main role. Indeed, if the first individual appears in a place where the growth is low or negative, it has a high probability to die with no descendants.

We denote the probability measure under which $\eta_0 = \delta_x$ by \mathbb{P}_{δ_x} .

Theorem 5.2. Let H be the principal eigenvalue of the elliptic operator $m\Delta_x + (b - d)$, with Neumann boundary conditions on \mathcal{X} , see (2.2). If $H > 0$, there exists a unique positive C^2 -solution ϕ^* to the elliptic equation

$$\begin{cases} 0 = m\Delta_x \phi^*(x) + (b(x) - d(x))\phi^*(x) - b(x)\phi^*(x)^2, & \forall x \in \mathcal{X}, \\ \partial_n \phi^*(x) = 0, & \forall x \in \partial\mathcal{X}, \end{cases} \quad (5.1)$$

and $\phi^*(x) = \lim_{t \rightarrow \infty} \mathbb{P}_{\delta_x}[\Upsilon_0 \geq t]$ for all $x \in \mathcal{X}$.

If $H \leq 0$, (5.1) has no non negative solution, we set $\phi^* \equiv 0$, and $\lim_{t \rightarrow \infty} \mathbb{P}_{\delta_x}[\Upsilon_0 \geq t] = 0$ for all $x \in \mathcal{X}$.

The second result of this part estimates the probability that the population size is of order K after a logarithmic time $\log K$. For all $\epsilon > 0$ and $K \in \mathbb{N}^*$, we set

$$\Upsilon_{\epsilon K} = \inf\{t \geq 0, \langle \eta_t, \mathbf{1} \rangle \geq \epsilon K\}.$$

Theorem 5.3. Let $\epsilon > 0$ and $(t_K)_{K>0}$ be a sequence of times such that $\lim_{K \rightarrow +\infty} t_K / \log(K) = +\infty$. Then for all $x \in \mathcal{X}$,

$$\lim_{K \rightarrow +\infty} \mathbb{P}_{\delta_x}[\Upsilon_{\epsilon K} < t_K] = \phi^*(x).$$

The end of this part is devoted to the two proofs.

Proof of Theorem 5.2. We first study the behavior of the probability $\mathbb{P}_{\delta_x}[\Upsilon_0 \leq t]$. We denote the time of the first event (birth or death) of the population by E_1 . The law of E_1 is given by $\mathbb{P}_{\delta_x}[E_1 \leq t] = \mathbb{E}_{\delta_x} \left[\int_0^t (d(X_s) + b(X_s)) e^{-\int_0^s (b(X_r) + d(X_r)) dr} ds \right]$. We set $I(s) := \int_0^s (b(X_r) + d(X_r)) dr$. Using the Markov property of η , we obtain

$$\begin{aligned} \mathbb{P}_{\delta_x}[\Upsilon_0 \leq t] &= \mathbb{E}_{\delta_x} \left[\mathbf{1}_{E_1 \leq t} \mathbf{1}_{\{M_{E_1}=0\}} + \mathbf{1}_{E_1 \leq t} \mathbf{1}_{\{M_{E_1}=2\}} \mathbb{E}_{2\delta_{X_{E_1}}} [\mathbf{1}_{M_{t-E_1}=0}] \right] \\ &= \mathbb{E}_x \left[\int_0^t \left(d(X_s) + b(X_s) \mathbb{P}_{\delta_{X_s}}[\Upsilon_0 \leq t-s]^2 \right) e^{-I(s)} ds \right], \end{aligned}$$

where X under \mathbb{P}_x is solution to (1.2) with the initial condition x and the diffusion coefficient m . Thus $g(x, t) = \mathbb{P}_{\delta_x}[\Upsilon_0 \leq t]$ satisfies for all $x \in \mathcal{X}$, and all $t > 0$,

$$\begin{cases} g(x, t) = \mathbb{E}_x \left[\int_0^t \left(d(X_s) + b(X_s) g(X_s, t-s)^2 \right) e^{-I(s)} ds \right], \quad \forall (x, t) \in \mathcal{X} \times \mathbb{R}^+, \\ g(x, 0) = 0, \quad \forall x \in \mathcal{X}. \end{cases} \quad (5.2)$$

Using Gronwall's Lemma for bounded functions, we deduce immediately that (5.2) has a unique bounded solution.

We will now show that there exists a unique C^2 -solution to

$$\begin{cases} \partial_t f(x, t) = m \Delta_x f(x, t) - (b(x) + d(x)) f(x, t) + d(x) + b(x) f(x, t)^2, \quad \forall (x, t) \in \mathcal{X} \times \mathbb{R}^+, \\ \partial_n f(x, t) = 0, \quad \forall (x, t) \in \partial \mathcal{X} \times \mathbb{R}^+ \\ f(x, 0) = 0, \quad \forall x \in \mathcal{X}, \end{cases} \quad (5.3)$$

such that $f \in C^{2,1}(\mathcal{X} \times \mathbb{R}^+)$, is positive and is smaller than 1 by using super- and sub-solutions arguments. Indeed, let $F(x, f) = -(b(x) + d(x))f + d(x) + b(x)f^2$. We easily see that $\underline{f} \equiv 0$ and $\bar{f} \equiv 1$ satisfy:

$$\begin{aligned} \begin{cases} \partial_t \underline{f} \leq m \Delta_x \underline{f} + F(x, \underline{f}), \quad \forall (x, t) \in \mathcal{X} \times \mathbb{R}^+, \\ \partial_t \bar{f} \geq m \Delta_x \bar{f} + F(x, \bar{f}), \quad \forall (x, t) \in \mathcal{X} \times \mathbb{R}^+, \\ \underline{f}(x, 0) \leq f(x, 0) \leq \bar{f}(x, 0), \quad x \in \mathcal{X}, \\ \partial_n \underline{f}(x, t) \leq 0 \leq \partial_n \bar{f}(x, t), \quad x \in \partial \mathcal{X}, t \in \mathbb{R}^+. \end{cases} \end{aligned}$$

That is, \underline{f} (resp. \bar{f}) is a sub-solution (resp. super-solution) to (5.3). Moreover, F and $\partial_f F$ belong to $C(\mathcal{X} \times \mathbb{R})$ and F is a Lipschitz function with respect to x by means of Assumptions 5.1. We apply Theorem 4 of Chapter III of [29] to deduce that (5.3) admits a solution $f \in C^{2,1}(\mathcal{X} \times \mathbb{R}^+)$ satisfying $0 \leq f \leq 1$. The uniqueness of the solution is a consequence of the maximum principle.

The next step is to use a Feynman-Kac formula to deduce that f is also a solution to (5.2). Indeed, for X solution to (1.2), let, for all $t \geq 0$, and $s \in [0, t]$,

$$H(s, X_s) = f(X_s, t-s) e^{-I(s)}.$$

Applying Itô Formula to $H(s, X_s)$, and using (1.2), (5.3) and the fact that $\partial_n f(x, t) = 0$ for all $x \in \partial\mathcal{X}$, we find for all $s \in [0, t]$,

$$\begin{aligned} H(s, X_s) &= H(0, X_0) - \int_0^s \left(d(X_\sigma) + b(X_\sigma)f(X_\sigma, t - \sigma)^2 \right) e^{-I(\sigma)} d\sigma \\ &\quad + \int_0^s \sqrt{2m}(\partial_x f(X_\sigma, t - \sigma)) e^{-I(\sigma)} dB_\sigma. \end{aligned} \quad (5.4)$$

We take the expectation of (5.4) for $s < t$. The expectation of the last term is equal to 0 as $(\int_0^r \partial_x f(X_\sigma, t - \sigma) e^{-I(\sigma)} dB_\sigma)_{r \in [0, s]}$ is a martingale. In addition, $\mathbb{E}_x[H(0, X_0)] = f(x, t)$. It stays to make s tend to t using the dominate convergence Theorem. As $\mathbb{E}_x[H(s, X_s)] \xrightarrow{s \rightarrow t} \mathbb{E}_x[H(t, X_t)] = 0$, we deduce that f is a solution to (5.2). (5.2) admits a unique bounded solution, thus both solutions are equal, i.e. $\mathbb{P}_{\delta_x}[\Upsilon_0 \leq t] = f(x, t)$.

Finally, we deduce the survival probability $\lim_{t \rightarrow +\infty} (1 - f(x, t))$ using results on Equation (5.3) described in [2] and in Theorems 9 and 11 of Chapter III in [29]. Indeed, they prove that if $H > 0$, there exists a unique positive solution ϕ^* to the elliptic equation (5.1), and that $\phi(t, x) = 1 - f(x, t) = \mathbb{P}_x(\Upsilon_0 > t)$ tends to $\phi^*(x)$ in $C^2(\mathcal{X})$ as $t \rightarrow +\infty$. Moreover, if $H \leq 0$, the unique solution to (5.1) is the zero function and $\phi(t, x) \rightarrow 0$ uniformly in \mathcal{X} as $t \rightarrow +\infty$. \square

Proof of Theorem 5.3. First, we split the studied probability into three parts :

$$\begin{aligned} \mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} < t_K) &= \mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} < t_K, \log \log(K) < \Upsilon_0 < +\infty) \\ &\quad + \mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} < t_K, \log \log(K) \geq \Upsilon_0) + \mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} < t_K, \Upsilon_0 = +\infty). \end{aligned} \quad (5.5)$$

Let us start with the first term of (5.5) :

$$\mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} < t_K, \log \log(K) < \Upsilon_0 < +\infty) \leq \mathbb{P}_{\delta_x}(\log \log(K) < \Upsilon_0 < +\infty) \xrightarrow{K \rightarrow +\infty} 0.$$

The second term of (5.5) will be treated using a comparison with a pure birth process. Let us consider a birth process with constant birth rate \bar{b} and started with only one individual. $\tilde{\Upsilon}_{\epsilon K}$ denote the first time when \tilde{N} , the population size of the process, is greater than ϵK .

$$\begin{aligned} \mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} < t_K, \log \log(K) \geq \Upsilon_0) &\leq \mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} \leq \log \log(K)) \\ &\leq \mathbb{P}_1(\tilde{\Upsilon}_{\epsilon K} \leq \log \log(K)) \\ &\leq \mathbb{P}_1(\tilde{N}_{\log \log K} \geq \epsilon K) \\ &\leq e^{\bar{b} \log \log K} / (\epsilon K) \xrightarrow{K \rightarrow +\infty} 0. \end{aligned}$$

There remains to deal with the third term in (5.5). Note that if $H \leq 0$, Theorem (5.2) implies that the third term is equal to zero, and the proof is done.

From this point forward, we assume that $H > 0$. Let h be a positive eigenvector of the

operator $m\Delta_x + (b-d)$. with Neumann boundary conditions on $\partial\mathcal{X}$ associated with the eigenvalue H . Thanks to Itô's Formula, we find

$$\langle \eta_t, e^{-Ht}h \rangle = \langle \eta_0, h \rangle + \int_0^t \langle \eta_s, e^{-Hs}m\Delta_x h + (b-d)he^{-Hs} - Hhe^{-Hs} \rangle ds + \int_0^t \sum_{i=1}^{M_s} \sqrt{2m} \nabla_x h(X_s^i) e^{-Hs} dB_s^i,$$

As $m\Delta_x h + (b-d)h = Hh$ and $\nabla_x h$ is bounded on \mathcal{X} , $(e^{-Ht}\langle \eta_t, h \rangle)_{t \geq 0}$ is a martingale. Moreover it is positive, so, it converges a.s. to a non-negative random variable that will be denoted by W . Obviously, $\{\Upsilon_0 < +\infty\} \subset \{W = 0\}$. Our aim is to prove that this is an a.s. equality. As done in the previous proof, we denote the time of the first event of the population by E_1 and we set $I(s) := \int_0^s (b(X_r) + d(X_r))dr$. Using the Markov property and the independence between individuals, we find an equation satisfied by $\mathbb{P}_{\delta_x}[W = 0]$:

$$\mathbb{P}_{\delta_x}[W = 0] = \mathbb{E}_x \left[\int_0^{+\infty} e^{-I(s)} \left(d(X_s) + b(X_s) \mathbb{P}_{\delta_{X_s}}[W = 0] \right) ds \right]. \quad (5.6)$$

Finally, $g(x) := \mathbb{P}_{\delta_x}[W > 0] = 1 - \mathbb{P}_{\delta_x}[W = 0]$ is solution to

$$g(x) = \mathbb{E}_x \left[\int_0^{+\infty} e^{-I(s)} b(X_s) g(X_s) (2 - g(X_s)) ds \right], \forall x \in \bar{\mathcal{X}}. \quad (5.7)$$

Let us show that there exists at most one non-zero solution with value in $[0, 1]$ to Equation (5.7). Let g_1 and g_2 be two such solutions. We define

$$\gamma = \sup\{\tilde{\gamma} > 0, g_1(x) - \tilde{\gamma}g_2(x) \geq 0, \forall x \in \bar{\mathcal{X}}\}.$$

Assume first that $\gamma < 1$.

$$g_1(x) - \gamma g_2(x) = \mathbb{E}_x \left[\int_0^{+\infty} e^{-I(s)} b(X_s) \left[2(g_1(X_s) - \gamma g_2(X_s)) - (g_1(X_s)^2 - \gamma g_2(X_s)^2) \right] ds \right]. \quad (5.8)$$

As $\gamma < 1$, $g_1^2 - \gamma g_2^2 \leq (g_1 - \gamma g_2)(g_1 + \gamma g_2)$ and we find

$$g_1(x) - \gamma g_2(x) \geq \mathbb{E}_x \left[\int_0^{+\infty} e^{-I(s)} b(X_s) [g_1(X_s) - \gamma g_2(X_s)] [2 - g_1(X_s) - \gamma g_2(X_s)] ds \right].$$

Moreover, by the definition of γ , there exists $x_0 \in \bar{\mathcal{X}}$ such that $g_1(x_0) - \gamma g_2(x_0) = 0$, so

$$0 = \mathbb{E}_{x_0} \left[\int_0^{+\infty} e^{-I(s)} b(X_s) [g_1(X_s) - \gamma g_2(X_s)] [2 - g_1(X_s) - \gamma g_2(X_s)] ds \right].$$

Thus for a.e. $s \in \mathbb{R}^+$, \mathbb{P}_{x_0} -a.s., $b(X_s)[g_1(X_s) - \gamma g_2(X_s)][2 - g_1(X_s) - \gamma g_2(X_s)] = 0$.

Let us note that for all $x \in \bar{\mathcal{X}}$, for $i = 1, 2$, (5.6) implies that

$$1 - g_i(x) \geq \mathbb{E}_x \left[\int_0^{+\infty} e^{-I(s)} d(X_s) ds \right] = \mathbb{P}_{\delta_x}[M_{E_1} = 0] > 0,$$

that is, for all $x \in \bar{\mathcal{X}}$, $2 - g_1(x) - \gamma g_2(x) > 0$. As b is positive, for a.e. $s \in \mathbb{R}^+$, \mathbb{P}_{x_0} -a.s. $g_1(X_s) - \gamma g_2(X_s) = 0$. In addition with the fact that $\gamma g_2^2 - g_1^2 = \gamma g_2^2(1 - \gamma) - (g_1 - \gamma g_2)(g_1 + \gamma g_2)$, (5.8) implies that,

$$\begin{aligned} 0 &= g_1(x_0) - \gamma g_2(x_0) \\ &= \mathbb{E}_{x_0} \left[\int_0^{+\infty} e^{-I(s)} b(X_s) \gamma g_2(X_s) [1 - \gamma] ds \right]. \end{aligned}$$

Using the same argument as before and that $1 - \gamma > 0$, we deduce that for a.e. $s \in \mathbb{R}^+$, \mathbb{P}_{x_0} -a.s., $g_2(X_s) = 0$. Moreover, under \mathbb{P}_{x_0} , the random variable X_s has a density with respect to Lebesgue measure, that is, for Lebesgue-a.a. $x \in \bar{\mathcal{X}}$, $g_2(x) = 0$. This is a contradiction with the fact that g_2 is a non-zero solution.

Finally if $\gamma \geq 1$, we define instead $\gamma' = \sup\{\tilde{\gamma} > 0, g_2(x) - \tilde{\gamma} g_1(x) \geq 0, \forall x \in \mathcal{X}\} < 1$ and we use symmetric arguments to reach a contradiction. Thus, there is at most one solution to (5.7) with values in $[0, 1]$.

The next step is to show that ϕ^* , which is solution to (5.1), is also solution to (5.7). Let us write $f^* = 1 - \phi^*$, which satisfies

$$\begin{cases} 0 = m \Delta_x f^* - (b + d) f^* + d + b(f^*)^2, & \text{on } \mathcal{X}, \\ \partial_n f^* = 0, & \text{on } \partial \mathcal{X}. \end{cases} \quad (5.9)$$

We apply Itô's Formula to $f^*(X_t) e^{-\int_0^t (b(X_r) + d(X_r)) dr}$. Then taking the expectation and using Equation (5.9), we deduce

$$f^*(x) = \mathbb{E}_x \left[\int_0^t e^{-I(s)} [d(X_s) + b(X_s) f^*(X_s)^2] ds \right] + \mathbb{E}_x \left[f^*(X_t) e^{-I(t)} \right].$$

Our aim is now to let t tend to infinity. Note that $I(t) \geq \underline{b}t$ for all $t \in \mathbb{R}^+$ and that f^* is bounded by 1. Hence, we use the dominated convergence Theorem to find those two convergences:

$$\begin{aligned} \left| \mathbb{E}_x \left[f^*(X_t) e^{-I(t)} \right] \right| &\leq \mathbb{E}_x \left[e^{-I(t)} \right] \xrightarrow{t \rightarrow +\infty} 0 \\ \left| \mathbb{E}_x \left[\int_t^{+\infty} e^{-I(s)} [d + b(f^*)^2](X_s) ds \right] \right| &\leq \mathbb{E}_x \left[\int_t^{+\infty} e^{-I(s)} (\bar{b} + \bar{d}) ds \right] \xrightarrow{t \rightarrow +\infty} 0. \end{aligned}$$

Thus, we make t tend to infinity and we find, for all $x \in \bar{\mathcal{X}}$,

$$f^*(x) = \mathbb{E}_x \left[\int_0^{+\infty} e^{-I(s)} [d(X_s) + b(X_s) f^*(X_s)^2] ds \right].$$

Since $\phi^* \equiv 1 - f^*$, we conclude that ϕ^* is a solution to (5.7). There exists at most one non-zero solution to (5.7), thus we have either $\mathbb{P}_{\delta_x}[W > 0] = \phi^*(x)$ for all $x \in \bar{\mathcal{X}}$, or $\mathbb{P}_{\delta_x}[W > 0] = 0$ for all $x \in \bar{\mathcal{X}}$. Using Itô's Formula, it is easy to check that in the case

$H > 0$, $(\langle \eta_t, e^{-Ht} h \rangle)_{t \geq 0}$ is bounded in L^2 . So this martingale is uniformly bounded and it converges in L^1 to W , hence $\mathbb{E}_{\delta_x}[W] = h(x) > 0$. Finally,

$$\mathbb{P}_{\delta_x}[W > 0] = \phi^*(x) = \mathbb{P}_{\delta_x}[\Upsilon_0 = +\infty] \Rightarrow \{W > 0\} = \{\Upsilon_0 = +\infty\} \text{ a.s.} \quad (5.10)$$

Thus, on $\{\Upsilon_0 = +\infty\}$,

$$\frac{\log(\epsilon K)}{\Upsilon_{\epsilon K}} \geq \frac{\log(\langle \eta_{\Upsilon_{\epsilon K}}, h e^{-H\Upsilon_{\epsilon K}} \rangle \cdot \|h\|_{\infty}^{-1} e^{H\Upsilon_{\epsilon K}})}{\Upsilon_{\epsilon K}} \xrightarrow{K \rightarrow +\infty} H > 0 \text{ a.s.},$$

as $\langle \eta_{\Upsilon_{\epsilon K}}, h \rangle \leq \epsilon K \|h\|_{\infty}$, $\Upsilon_{\epsilon K} \rightarrow +\infty$ when K tends to infinity and $W > 0$. Hence,

$$\lim_{K \rightarrow +\infty} \frac{\Upsilon_{\epsilon K}}{\log(\epsilon K)} < +\infty \text{ and } \lim_{K \rightarrow +\infty} \frac{t_K}{\log(\epsilon K)} = +\infty,$$

and so, the third term in (5.5) satisfies

$$\mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} < t_K, \Upsilon_0 = +\infty) = \mathbb{P}_{\delta_x} \left(\frac{\Upsilon_{\epsilon K}}{\log(\epsilon K)} < \frac{t_K}{\log(\epsilon K)}, \Upsilon_0 = +\infty \right) \xrightarrow{K \rightarrow \infty} \mathbb{P}_{\delta_x}(\Upsilon_0 = +\infty). \quad (5.11)$$

Finally, we have shown that the two first terms in (5.5) tend to 0, so using additionally (5.11)

$$\lim_{K \rightarrow +\infty} \mathbb{P}_{\delta_x}(\Upsilon_{\epsilon K} < t_K) = \mathbb{P}_{\delta_x}(\Upsilon_0 = +\infty) = \phi^*(x).$$

That ends the proof for $H > 0$. □

6 Proof of Theorem 2.5

This section is devoted to prove Theorem 2.5. The structure of the proof of Theorem 2.5 is similar to the one of [3], thus, we do not repeat all the details but only the points that are different.

The first proposition concerns the behavior of the first time of a mutation S_1^K , when the initial state is a monomorphic population.

Proposition 6.1. *Suppose that Assumptions 1.1 and (2.4) hold. Let $u \in \mathcal{U}$ and \mathcal{C}^u a compact subset of $M_F(\mathcal{X} \times \{u\})$ such that $0 \notin \mathcal{C}^u$. If $\nu_0^K \in \mathcal{C}^u \cup M_F^K(\mathcal{X})$, then*

- *for any $\gamma > 0$, $\lim_{K \rightarrow +\infty} \mathbb{P}_{\nu_0^K} \left(S_1^K > \log K, \sup_{t \in [\log K, S_1^K]} \mathcal{W}_1(\nu_t^K, \bar{\xi}^u \delta_u) \geq \gamma \right) = 0$.*
- *Furthermore, $\lim_{K \rightarrow +\infty} \mathbb{P}_{\nu_0^K} \left(S_1^K > \frac{t}{Kq_K} \right) = \exp \left(-t \int_{\mathcal{X}} b^u(x) p^u(x) \bar{\xi}^u(dx) \right)$.*

Proposition 6.1 can be proved using similar arguments as those of the proof of Lemma 2 in [3]. It is a consequence of the following lemma.

Lemma 6.2. *For any $\alpha > 0$, there exists $T_\alpha > 0$ such that for any $\xi_0 \in \mathcal{C}^u$, for any $t \geq T_\alpha$, $\mathcal{W}_1(\xi_t, \bar{\xi}^u) < \alpha$, where $(\xi_t)_{t \geq 0}$ is the solution to Equation (2.1) with initial state ξ_0 .*

Proof. On the one hand, Theorem 1.4 in [22] implies that the density of $\bar{\xi}^u$ is a stable monomorphic equilibrium for the L^2 -distance. Using a proof in three steps as that of Proposition 4.2, we prove a \mathcal{W}_1 -stability : there exists α' such that for any $\xi_0 \in B(\bar{\xi}^u, \alpha')$ and for any $t \geq 0$, $\mathcal{W}_1(\xi_t, \bar{\xi}^u) < \alpha$.

On the other hand, for any $\xi_0 \in \mathcal{C}^u$, ξ_t converges towards $\bar{\xi}^u$. There exists T_{ξ_0} such that $\mathcal{W}_1(\xi_{T_{\xi_0}}, \bar{\xi}^u) < \alpha'/2$. Using Lemma 3.5 and arguments similar to (4.5) and (4.6), we show that for any $t \geq 0$, any $\zeta_0 \in M_F(\mathcal{X})$,

$$\sup_{r \in [0, t]} \mathcal{W}_1(\zeta_r, \xi_r) \leq C(t) \mathcal{W}_1(\zeta_0, \xi_0),$$

where $(\zeta_t)_{t \geq 0}$ is the solution to (2.1) with initial state $\zeta_0 \in M_F(\mathcal{X} \times \{u\})$. Consequently, there exists $\delta_{\xi_0} > 0$ such that for any $\zeta_0 \in B(\xi_0, \delta_{\xi_0})$, $\mathcal{W}_1(\zeta_{T_{\xi_0}}, \xi_{T_{\xi_0}}) < \alpha'/2$. Thus, for any ζ_0 such that $\mathcal{W}_1(\zeta_0, \xi_0) < \delta_{\xi_0}$, for any $t \geq T_{\xi_0}$, $\mathcal{W}_1(\zeta_t, \bar{\xi}^u) < \alpha$.

Finally, as \mathcal{C}^u is a compact set, there exists a finite number of balls such that $\mathcal{C}^u \subset \cup_{i=1}^n B(\xi_0^i, \delta_{\xi_0^i})$. Defining $T_\alpha = \max_{i=1..n} T_{\xi_0^i}$, we deduce the lemma. \square

Proof of Proposition 6.1. First, remark that the first probability of Proposition 6.1 is non-increasing with γ . Thus, it is sufficient to prove the property for any small $\gamma > 0$. Let us assume that γ satisfies the assumptions of Theorem 4.1. Theorem 4.1 in the monomorphic case implies that : there exist $\gamma' > 0$, $V > 0$, such that

$$\sup_{\nu_0^K \in B(\bar{\xi}^u, \gamma') \cap M_F^K} \mathbb{P}(R_\gamma^K \leq S_1^K \wedge e^{KV}) \xrightarrow{K \rightarrow \infty} 0, \quad (6.1)$$

R_γ^K is defined by (4.1). We set $2\alpha = \gamma'$, then Lemma 6.2 and Theorem 3.1 imply that

$$\sup_{\nu_0^K \in \mathcal{C}^u} \mathbb{P}(\mathcal{W}_1(\nu_{T_\alpha}^K, \bar{\xi}^u) \geq 2\alpha) \xrightarrow{K \rightarrow +\infty} 0. \quad (6.2)$$

Using the Markov property, we deduce if K is sufficiently large such that $\log(K) > T_\alpha$,

$$\begin{aligned} & \mathbb{P}_{\nu_0^K} \left(S_1^K > \log K, \sup_{t \in [\log K, S_1^K]} \mathcal{W}_1(\nu_t^K, \bar{\xi}^u) \geq \gamma \right) \\ & \leq \mathbb{P}_{\nu_0^K}(\mathcal{W}_1(\nu_{T_\alpha}^K, \bar{\xi}^u) \geq \gamma') + \mathbb{E}_{\nu_0^K} \left[\mathbb{1}_{\{S_1^K \geq \log(K), \mathcal{W}_1(\nu_{T_\alpha}^K, \bar{\xi}^u) < \gamma'\}} \sup_{t \in [T_\alpha, S_1^K + T_\alpha]} \mathcal{W}_1(\nu_t^K, \bar{\xi}^u) \geq \gamma \right] \\ & \leq \mathbb{P}_{\nu_0^K}(\mathcal{W}_1(\nu_{T_\alpha}^K, \bar{\xi}^u) \geq \gamma') + \mathbb{E}_{\nu_0^K} \left[\mathbb{1}_{\{S_1^K \geq \log(K), \mathcal{W}_1(\nu_{T_\alpha}^K, \bar{\xi}^u) < \gamma'\}} \mathbb{P}_{\nu_{T_\alpha}^K}(R_\gamma^K \leq S_1^K) \right] \\ & \leq \mathbb{P}_{\nu_0^K}(\mathcal{W}_1(\nu_{T_\alpha}^K, \bar{\xi}^u) \geq 2\alpha) + \sup_{\nu_0 \in B(\bar{\xi}^u, \gamma')} \left[\mathbb{P}_{\nu_0}(R_\gamma^K \leq S_1^K \wedge e^{KV}) + \mathbb{P}_{\nu_0}(e^{KV} \leq R_\gamma^K \leq S_1^K) \right]. \end{aligned}$$

The first term and the second term tend to 0 when K tends to $+\infty$ according to (6.2) and (6.1) respectively. There remains the third term. On $\{t \leq R_\gamma^K \leq S_1^K\}$, the number of mutations \mathcal{M}_t is stochastically bounded from below by a Poisson process with parameter $Kq_K(\langle b^u p, \bar{\xi}^u \rangle - \gamma \|b^u p\|_{Lip})$ which is positive if γ is small enough. We conclude with the fact that $\mathbb{P}(e^{KV} \leq R_\gamma^K \leq S_1^K) \leq \mathbb{P}(\mathcal{M}_{e^{KV}} = 0) \xrightarrow{K \rightarrow +\infty} 0$, under Assumption (2.4). That ends the proof of the first point.

The second point of Proposition 6.1 is easily deduced from this first point and Lemma 2 in [3]. \square

The second proposition studies the process with a dimorphic initial state. Let us define :

- θ_0 is the first time when the population becomes monomorphic again,
- V_0 is the phenotypic trait of the population at time θ_0 .

Proposition 6.3. *Suppose Assumptions 1.1, 2.3 and (2.4) hold, and that the initial state $\nu_0^K \in M_F(\mathcal{X} \times \{u, v\})$ is such that $\nu_0^{K,u}$ converges weakly to $\bar{\xi}^u$ in $M_F(\mathcal{X})$ and $\nu_0^{K,v} = \frac{\delta_{x_0}}{K}$. Then,*

$$\lim_{K \rightarrow \infty} \mathbb{P}(V_0 = v) = 1 - \lim_{K \rightarrow \infty} \mathbb{P}(V_0 = u) = \phi^{vu}(x_0).$$

Moreover,

$$\forall \eta > 0, \lim_{K \rightarrow \infty} \mathbb{P}\left(\theta_0 \leq S_1^K \wedge \frac{\eta}{Kq_K}\right) = 1, \text{ and } \forall \gamma > 0, \lim_{K \rightarrow \infty} \mathbb{P}(\mathcal{W}_1(\nu_{\theta_0}^K, \bar{\xi}^{V_0}) < \gamma) = 1.$$

Proof. We set $\gamma > 0$ small enough to use Theorem 4.1 : there exist γ', ϵ, V such that

$$\sup_{\nu_0^{K,u} \in B(\bar{\xi}^u, \gamma'), \nu_0^{K,v} \in B(0, \epsilon)} \mathbb{P}_{\nu_0^K}(R_\gamma^K \geq T_\epsilon^K \wedge S_1^K \wedge e^{KV}) \xrightarrow{K \rightarrow \infty} 0,$$

where R_γ^K, T_ϵ^K have been defined by (4.1), (4.3).

Let assume that $\epsilon \leq \gamma$ and that K is large enough such that $q_K \leq \gamma$. Then on $\{t \leq R_\gamma^K \wedge T_\epsilon^K \wedge S_1^K\}$, the process $(\nu_t^{K,v})_{t \geq 0}$ is stochastically bounded:

$$\frac{1}{K} Z^{inf} \preceq \nu^{K,v} \preceq \frac{1}{K} Z^{sup},$$

Z^{inf}, Z^{sup} are two branching diffusion processes starting with one individual at location x_0 , their birth rates are respectively $b^v(x)(1 - \gamma)$ and $b^v(x)$, and their death rates are $d^v(x) + c^{vu} \cdot \bar{\xi}^u + 2\bar{c}\gamma$ and $d^v(x) + c^{vu} \cdot \bar{\xi}^u - \bar{c}\gamma$.

Let us set

$$\Upsilon_{K\epsilon}^{sup} = \inf\{t \geq 0, \langle Z_t^{sup}, \mathbb{1} \rangle \geq K\epsilon\},$$

$$\Upsilon_0^{sup} = \inf\{t \geq 0, \langle Z_t^{sup}, \mathbb{1} \rangle = 0\},$$

and respectively $\Upsilon_{K\epsilon}^{inf}, \Upsilon_0^{inf}$ associated with Z^{inf} .

Using same kind of computations as in Lemma 3 in [3], we deduce that

$$\begin{aligned} & \mathbb{P}_{\nu_0^K} \left(\theta_0 \leq S_1^K \wedge \frac{\eta}{q_K K}, V_0 = u, \mathcal{W}_1(\nu_{\theta_0}^K, \bar{\xi}^u) < \gamma \right) \\ & \geq \mathbb{P}_{\delta_{x_0}} \left(\Upsilon_0^{sup} \leq \frac{\eta}{q_K K} \wedge \Upsilon_{K\epsilon}^{sup} \right) - \mathbb{P}_{\nu_0^K} \left(\frac{\eta}{q_K K} \geq S_1^K \right) - \mathbb{P}_{\nu_0^K} \left(\frac{\eta}{q_K K} \wedge S_1^K \wedge T_\epsilon^K \geq R_\gamma^K \right). \end{aligned} \quad (6.3)$$

Theorem 4.1 implies that the last term tends to 0 under Assumption (2.4). The second term tends also to 0 as the number of individuals is stochastically bounded from above by a birth and death process with birth rate \bar{b} and competition rate \underline{c} , thus Lemma 2 in [3] implies that for any $\delta > 0$, there exists $\eta > 0$ such that $\limsup_{K \rightarrow +\infty} \mathbb{P}(S_1^K \leq \frac{\eta}{K q_K}) \leq \delta$. Thus the main difficulty is to evaluate the first term. Theorem 5.2 implies that

$$\mathbb{P}_{\delta_{x_0}} \left(\Upsilon_0^{sup} \leq \frac{\eta}{q_K K} \wedge \Upsilon_{K\epsilon}^{sup} \right) \xrightarrow{K \rightarrow +\infty} \mathbb{P}_{\delta_{x_0}}(\Upsilon_0^{sup} < +\infty) = 1 - \phi^{\gamma, vu}(x_0), \quad (6.4)$$

where $\phi^{\gamma, vu}$ is the solution to the following elliptic equation on \mathcal{X} with Neumann boundary condition

$$m^v \Delta_x \phi^{\gamma, vu}(x) + [b^v(x) - d^v(x) - c^{vu} \cdot \bar{\xi}^u + \bar{c}\gamma] \phi^{\gamma, vu}(x) - b^v(x) \phi^{\gamma, vu}(x)^2 = 0.$$

Let us show that this solution is close to ϕ^{vu} . Theorem 5.2 implies that $\phi^{\gamma, vu}$ is positive if and only if

$$H^v - c^{vu} \cdot \bar{\xi}^u + \bar{c}\gamma = H^v - \frac{H^u \kappa^{vu}}{\kappa^{uu}} + \bar{c}\gamma > 0. \quad (6.5)$$

First case: $H^v \kappa^{uu} - H^u \kappa^{vu} < 0$ (Point 1 in Assumption 2.3)

We can find γ small enough such that (6.5) is not satisfied, thus $\phi^{\gamma, vu} \equiv \phi^{vu} \equiv 0$. In addition with (6.3) and (6.4), we deduce that

$$\mathbb{P}_{\nu_0^K} \left(\theta_0 \leq S_1^K \wedge \frac{\eta}{q_K K}, V_0 = u, \mathcal{W}_1(\nu_{\theta_0}^K, \bar{\xi}^u) < \gamma \right) \xrightarrow{K \rightarrow +\infty} 1.$$

Proposition 6.3 is proved for this case.

Second case: $H^v \kappa^{uu} - H^u \kappa^{vu} > 0$ (Point 2 in Assumption 2.3)

Hence (6.5) is satisfied for all $\gamma > 0$. Let C be $\frac{\bar{c}\gamma}{\inf_{y \in \mathcal{X}} b^v(y) \phi^{vu}(y)}$ and set $\mathcal{L}^\gamma(f) = m^v \Delta_x f + (b^v(x) - d^v(x) - c^{vu} \cdot \bar{\xi}^u + \bar{c}\gamma)f - b^v(x)f^2$. We have

$$\begin{aligned} \mathcal{L}^\gamma(\phi^{vu}) &= \bar{c}\gamma \phi^{vu} \geq 0, \\ \mathcal{L}^\gamma((1+C)\phi^{vu}) &= (1+C)\phi^{vu}[\bar{c}\gamma - Cb^v \phi^{vu}] \leq 0. \end{aligned}$$

As $\phi^{\gamma, vu}$ is the unique solution to $\mathcal{L}^\gamma(f) = 0$, we deduce the following inequalities from a comparison theorem (see for example Theorem III.5 in [29]) : for any $x \in \mathcal{X}$,

$$(1+C)\phi^{vu}(x) \geq \phi^{\gamma, vu}(x) \geq \phi^{vu}(x). \quad (6.6)$$

We split the end of the proof into three steps regarding as the proof of Lemma 3 in [3]. Let us fix $\alpha > 0$. (6.3), (6.4) and (6.6) imply that if K is large enough,

$$\mathbb{P}_{\nu_0^K} \left(T_0^K \leq \frac{\eta}{q_K K} \wedge S_1^K \wedge R_\gamma^K \wedge T_\epsilon^K \right) \geq 1 - \phi^{vu}(x_0) - \alpha.$$

On the other hand, if K is large enough :

$$\begin{aligned} & \mathbb{P}_{\nu_0^K} \left(T_\epsilon^K \leq \frac{\eta}{q_K K} \wedge S_1^K \wedge R_\gamma^K \right) \\ & \geq \mathbb{P}_{\delta_{x_0}} \left(T_{K\epsilon}^{inf} \leq \frac{\eta}{q_K K} \wedge T_0^{inf} \right) - \mathbb{P}_{\nu_0^K} \left(\frac{\eta}{q_K K} \geq S_1^K \right) - \mathbb{P}_{\nu_0^K} \left(\frac{\eta}{q_K K} \wedge S_1^K \wedge T_\epsilon^K \geq R_\gamma^K \right), \\ & \geq \phi^{vu}(x_0) - \alpha. \end{aligned}$$

Thus, the v -process $\nu^{K,v}$ reaches a non-negligible size $K\epsilon$, before $\frac{\eta}{q_K K} \wedge S_1^K \wedge R_\gamma^K$, with a probability that tends to $\phi^{vu}(x_0)$.

Once the mutant population has reached a non-negligible size, we can compare the stochastic process and the deterministic limiting process. Under point 2 in Assumption 2.3, there exist $T > 0$ and $\gamma_2 > 0$ such that for any $\xi_0 \in adh(B(\bar{\xi}^u, \gamma) \times (B(0, 2\epsilon) \setminus B(0, \epsilon)))$, for any $t \geq T$,

$$\xi_t \in B(0, \gamma_2) \times B(\bar{\xi}^v, \gamma_2),$$

where ξ is the solution to (2.1) with a dimorphic initial state $\xi_0 \in M_F(\mathcal{X} \times \{u, v\})$. This can be proved using similar arguments than those of the proof of Lemma 6.2.

Moreover, using Theorem 3.1 and Proposition 3.4,

$$\sup_{\nu_0^K} \mathbb{P}_{\nu_0^K} \left(\sup_{t \in [0, T]} \mathcal{W}_1(\nu_t^K, \xi_{t, \nu_0^K}) < \gamma_2 \right) \xrightarrow{K \rightarrow +\infty} 0,$$

with ξ_{t, ν_0^K} the solution to (2.1) with initial state ν_0^K . The two previous results and the Markov property imply that, if K is large enough,

$$\begin{aligned} & \mathbb{P}_{\nu_0^K} \left(T_\epsilon^K \leq \frac{\eta}{q_K K} \wedge S_1^K \wedge R_\gamma^K, S_1^K \geq \frac{\eta}{q_K K} + T, \nu_{T_\epsilon^K + T}^{K,u} \in B(0, 2\gamma_2), \nu_{T_\epsilon^K + T}^{K,v} \in B(\bar{\xi}^v, 2\gamma_2) \right) \\ & \geq \phi^{vu}(x_0) - 2\alpha. \end{aligned}$$

Finally, we use the Markov property at time $T_\epsilon^K + T$ and we conclude as in Lemma 3 in [3]. If γ_2 is sufficiently small, we prove that, with a probability that tends to 1, after time $T_\epsilon^K + T$, the u -population process $\nu^{K,u}$ will become extinct before its size reaches the threshold $\sqrt{\gamma_2}$ and before the v -process $\nu^{K,v}$ moves away from a neighborhood of the equilibrium $\bar{\xi}^v$.

That concludes the proof of Proposition 6.3 for the second case. \square

Theorem 2.5 is deduced from Propositions 6.1 and 6.3 in a similar way to Theorem 1 in [3] by using the transition probabilities of the jump process $(\Lambda_t)_{t \geq 0}$.

7 Numerics

To end this paper, let us illustrate Theorem 2.5 with a numerical example. We consider here a set of parameters similar to the one in [11] and [22]. The location space is $\mathcal{X} = (0, 1)$ and the trait space is $\mathcal{U} = [0, 1]$. For all trait u , we consider that the growth rate is maximal for the location $x = u$. For instance, the location space state can represent a variation of resources, as seed size for some birds, and so two populations with two different traits are not best-adapted to same resources. For some bird species, a gradual variation of seed size can determine a gradual variation of beak size [19]. Moreover, the maximum value of the growth rate function is the same for all traits but when u increases, the birth rate function goes faster to 0, as follows :

$$b(x, u) = \max\{4 - 160 \cdot u(x - u)^2, 0\}, \quad d(x, u) = 1.$$

That is, the birds with a small trait value are more generalists than the ones with a large trait value : they are adapted to a larger set of seed size [17]. All individuals move with the same diffusion coefficient $m^u = 0.003$. The competition kernel is a constant $c = 10$ and the mutation kernel $k(x, u, w)$ is the probability density of a Gaussian random variable with mean u and standard deviation 0.05 conditioned on staying in \mathcal{U} . The initial population is composed of K individuals at location $x = 0.5$ and with trait $u = 0.6$.

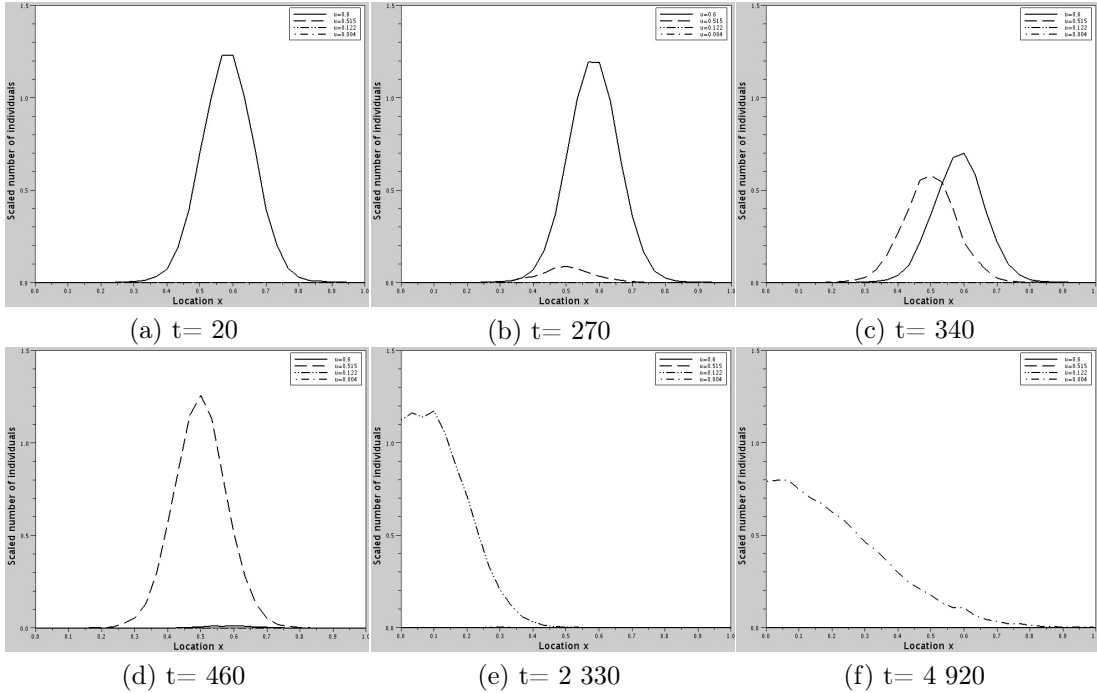


Figure 1: Simulations with $K = 100000$, $q_K = 10^{-5}$.

In Figure 1, we observe the evolution of the trait associated with a change of spatial niches and spatial patterns over time. After a short time (a), the spatial distribution of

the monomorphic population with trait $u = 0.6$ stabilizes. Then, in Figures (b), (c) and (d), we observe a phenomenon of invasion and replacement : some individuals with trait $u = 0.515$ appear, invade and finally replace the previous population with trait $u = 0.6$. Remark the change of spatial niche, see Figure (c). The locations of the population with trait $u = 0.515$ are slightly smaller than the one of trait $u = 0.6$. Three other phenomena of invasion and replacement with a displacement of the spatial niche are detected until the time $t = 2330$ (Figure (e)). In a second phase, the population evolve to become more and more generalists (Figure (f)) : the length of the spatial niche is increasing at each event of invasion and replacement.

The simulations are computed using the algorithm presented in Parts 3 and 6 in Champagnat, Méléard [5]. It is an iterative construction, which gives an effective algorithm of the process. The diffusion motion is simulated using an Euler scheme for reflected diffusion process.

Acknowledgments : I would like to thank S. Méléard for her guidance during my work. I also thank C. Léonard, G. Raoul, C. Tran and A. Veber for their help. I acknowledge partial support by the “Chaire Modélisation Mathématique et Biodiversité” of VEOLIA - École Polytechnique - MNHN - F.X.

References

- [1] A. Arnold, L. Desvillettes, and C. Prévost. Existence of nontrivial steady states for populations structured with respect to space and a continuous trait. *Commun. Pure Appl. Anal.*, 11(1):83–96, 2012.
- [2] H. Berestycki, F. Hamel, and L. Roques. Analysis of periodically fragmented environment model: I-species persistence. *J. Math. Biol.*, 51:75–113, 2005.
- [3] N. Champagnat. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic processes and their applications*, 116(8):1127–1160, 2006.
- [4] N. Champagnat, P.E. Jabin, and S. Méléard. Adaptation in a stochastic multi-resources chemostat model. *Journal de Mathématiques Pures et Appliquées*, 101(6):755–788, 2014.
- [5] N. Champagnat and S. Méléard. Invasion and adaptative evolution for individual-based spatially structured populations. *J. Math. Biol.*, 55(2):147–188, 2007.
- [6] M. Costa, C. Hauzy, N. Loeuille, and S. Méléard. Stochastic eco-evolutionary model of a prey-predator community. *arXiv preprint arXiv:1407.3069*, 2014.
- [7] J. Coville. Convergence to equilibrium for positive solutions of some mutation-selection model. *arXiv preprint arXiv:1308.6471*, 2013.
- [8] D. Dawson and J. Gärtner. Large deviations from the mc-kean-vlasov limit for weakly interacting diffusions. *Stochastics*, 20:247–308, 1987.
- [9] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Stochastic Modelling and Applied Probability, 1998.
- [10] Desvillettes, Ferrière, and Prévost. Infinite dimensional reaction-diffusion for population dynamics. *preprint*, 4(3):529–605, 2004.

- [11] M. Doebeli and U. Dieckmann. Speciation along environmental gradients. *Nature*, 421:259–263, 2003.
- [12] R. Durrett and S. Levin. Spatial aspects of interspecific competition. *Theoretical population biology*, 53(1):30–43, 1998.
- [13] J.A. Endler. *Geographic variation, speciation, and clines*. Number 10. Princeton University Press, 1977.
- [14] J. Fontbona. Uniqueness for a weak nonlinear evolution equation and large deviations for diffusing particles with electrostatic repulsion. *Stochastic processes and their applications*, 112(1):119–144, 2004.
- [15] M. Freidlin and A. Wentzell. *Random Perturbations*. Springer, 1984.
- [16] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, 1964.
- [17] D. Futuyma and G. Moreno. The evolution of ecological specialization. *Ann. Rev. Ecol. Syst.*, 19:207–233, 1988.
- [18] C. Graham and S. Méléard. An upper bound of large deviations for a generalized star-shaped loss network. *Markov Processes and Related Fields*, 3(2):199–224, 1997.
- [19] P. Grant and B. Grant. Unpredictable evolution in a 30-year study of darwin’s finches. *Science*, 296:707–711, 2002.
- [20] R. Kassen. The experimental evolution of specialists, generalists, and the maintenance of diversity. *J. Evol. Biol.*, 15:173–190, 2002.
- [21] O. Leimar, M. Doebeli, and U. Dieckmann. Evolution of phenotypic clusters through competition and local adaptation along an environmental gradient. *Evolution*, 62(4):807–822, 2008.
- [22] H. Leman, S. Méléard, and S. Mirrahimi. Influence of a spatial structure on the long time behavior of a competitive lotka-volterra type system. *Discrete Contin. Dyn. Syst. Ser. B*, pages 469–493, 2015.
- [23] C. Léonard. Convex conjugates of integral functionals. *Acta Math. Hungar.*, 93(4):253–280, 2001.
- [24] C. Léonard. Minimizers of energy functionals. *Acta Math. Hungar.*, 93(4):281–325, 2001.
- [25] C. Léonard. On large deviations for particle systems associated with spatially homogeneous boltzmann type equations. *Prob.Th.Rel.Fields*, 101(1):1–44, 1995.
- [26] J.A.J. Metz, S. Geritz, G. Meszéna, F. Jacobs, and J.S. Van Heerwaarden. Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. *Stochastic and spatial structures of dynamical systems*, 45:183–231, 1996.
- [27] J. Polechová and N.H. Barton. Speciation through competition: a critical review. *Evolution*, 59(6):1194–1210, 2005.
- [28] M.M. Rao and Z.D. Ren. *Applications of Orlicz spaces*, volume 250. CRC Press, 2002.
- [29] L. Roques. *Modèles de réaction-diffusion pour l’écologie spatiale*. Editions Quae, 2013.
- [30] D. Tilman and P.M. Kareiva. *Spatial ecology: the role of space in population dynamics and interspecific interactions*, volume 30. Princeton University Press, 1997.

- [31] V. Tran. Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM*, 12:345–386, 2008.
- [32] F. Wang and L. Yan. Gradient estimate on convex domains and applications. *Proceedings of the American Mathematical Society*, 141(3):1067–1081, 2013.